



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 96572**

**TO: Gary Nickol**  
**Location: cm-1/8d17/8e12**  
**Art Unit: 1642**  
**Friday, June 20, 2003**

**Case Serial Number: 513365**

**From: Mona Smith**  
**Location: Biotech-Chem Library**  
**CM1-6A01**  
**Phone: 308-3278**

**mona.smith@uspto.gov**

### **Search Notes**

Thank you for using STIC services  
Feel free to contact me if you have any questions.

Mona Smith  
308-3278



STIC-Biotech/ChemLib

96592

From: Nickol, Gary  
Sent: Friday, June 13, 2003 3:56 PM  
To: STIC-Biotech/ChemLib  
Subject: 09/513365

Please search the following:

- 1) DNA of SEQ ID NO:2
- 2) DNA encoding the amino acids of SEQ ID NO:1

Thanks,

Gary Nickol  
CM1, AU:1642  
Room 8D17, Mailbox 8E12  
(703) 305-7143

RECEIVED  
JUN 13 2003  
(STIC)

Searcher: M. Smith  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: 6/16/03  
Date Completed: 6/20/03  
Searcher Prep/Review: 4  
Clerical: 5  
Online time: 5

TYPE OF SEARCH:  
NA Sequences: 2  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 20, 2003, 02:40:48 ; Search time 1428 seconds

(without alignments)  
3175.586 Million cell updates/sec

Title: US-09-513-365A-1

Perfect score: 1481

Sequence: 1 MLCGQQQQQLYSSAALLTGER.....DNKTKMDKSTKTKKKDRSR 280

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=rlp  
-Q/cgnt2\_1/USFTO.spool/US09513365/runat\_16062003\_124743\_16453/app\_query.fasta.1.455  
-DB=EST-QPWT-fastap -SUFFIX=1st -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS-bits -START=1 -END=1 -MATRIX-BLOSUM62 -TRANS-human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USRR=US09513365 -CGN\_1\_1\_2874.grunat.16062003\_124743\_16453 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEPODTRY -NBS\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estlm:\*  
5: em\_estlov:\*  
6: em\_estlpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estlm:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrc:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1395.5	94.2	1464	11 AK012716	AK012716 Mus muscu
2	1296	87.5	912	14 BQ277444	BQ277444 AGNCOURT
3	1096.5	74.0	919	12 BG211544	BG211544 RST31111
4	1094	73.9	793	12 BE796780	BE796780 601587557
5	1063	71.8	778	13 B1548536	B1548536 603191255
6	1024	69.1	735	14 BM982877	BM982877 UT-CF-EN1
7	1006.5	68.0	798	12 BG184056	BG184056 RST2972 A
8	961	64.9	664	14 BM719460	BM719460 UT-E-BJ1
9	935	63.1	708	13 B1460319	B1460319 603201967
10	920	62.1	675	14 BM682789	BM682789 UT-E-BJ1
11	902	60.9	891	13 B1091047	B1091047 602855235
12	897	60.6	541	12 BF114684	BF114684 7166C08.X
13	889	60.0	679	10 AV647917	AV647917 AV647917
14	883	59.6	564	12 BE839427	BE839427 RC3-FN014
15	874	59.0	538	12 BF523624	BF523624 UT-R-C1-1
16	864	58.3	653	13 B1495873	B1495873 df121d04
17	856	57.8	522	13 B1495874	B1495874 df121d04
18	852	57.5	566	12 BE839460	BE839460 RC3-FN014
19	849	57.3	720	9 AA797495	AA797495 vW28B01.F
20	847	57.2	703	13 B1546850	B1546850 603189754
21	830	56.0	531	9 AA204672	AA204672 zrf8C06.F
22	829	56.0	525	9 A1885642	A1885642 w159h01.x
23	814	55.0	616	9 A1673816	A1673816 A1673816
24	805	54.4	547	9 A1186701	A1186701 qe82d12.x
25	798	53.9	482	12 BE831862	BE831862 RC6-WT006
26	780	52.7	692	9 A1633536	A1633536 B633536
27	779.5	52.6	830	12 BG189892	BG189892 RST8946 A
28	775	52.3	601	14 BQ206861	BQ206861 UT-M-BM1
29	766	51.7	844	12 BG862842	BG862842 602797509
30	762	51.5	975	14 BQ934291	BQ934291 ACENCOURT
31	760	51.3	513	9 A1857399	A1857399 w155f09.x
32	745	50.3	537	9 A1458721	A1458721 tk13h10.x
33	741	50.0	574	10 A1674008	A1674008 ba59f06.x
34	730	49.3	873	14 BQ430264	BQ430264 ACENCOURT
35	728	49.2	460	10 A1674790	A1674790 ba59f06.y
36	724	48.9	561	13 B1098630	B1098630 B1098630
37	709	47.9	682	9 AA746304	AA746304 oa36d03.F
38	700.5	47.3	609	10 AW342872	AW342872 IJ80d12.Y
39	697	47.1	475	10 BE653650	BE653650 UT-M-AH1
40	692	46.7	477	12 BF513151	BF513151 UT-H-BM1
41	690	46.6	555	14 BQ033210	BQ033210 UT-1-CFO-
42	688	46.5	502	12 BF732679	BF732679 nae05d11
43	682	46.0	520	13 B1676689	B1676689 iCS3d06.x
44	681.5	46.0	442	9 A1118087	A1118087 r2225a08
45	681	46.0	466	9 A1818729	A1818729 wk91d10.x

#### ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	1464 bp	mrna	linear	HTC 19-JAN-2002
AK012716	AK012716	Mus musculus	10, 11 days embryo	whole body	cdna, RIKEN full-length	
ACCESSION	AK012716	sequence.				
VERSION	AK012716.1	GI:12849649				
KEYWORDS	HTC; CAP trapper.					
SOURCE	Mus musculus (strain:C57BL/6J) 10, 11 days embryo	cdna to mRNA,				
ORGANISM	Mus musculus	clone:11b-RIKEN full-length				
		enriched mouse				
		cdna library				
		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				

1 Carninci, P. and Hayashizaki, Y.  
High efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Kono, H., Akiyama, J., Nishi, K., Kitsumai, T., Teshiro, H., Itoh, M.,  
Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwaka, E., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system—384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL MEDLINE  
PUBMED  
20530913  
11076861

4 Kawai, T., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
Aizawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S.,  
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamada, A.,  
Saito, T., Okazaki, Y., Gojodori, T., Bono, H., Kasukawa, T., Saito, R.,  
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,  
Fischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiya, H.,  
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,  
Quackenbush, J., Schriml, L. M., Stabile, F., Suzuki, R., Tomita, M.,  
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
Baldarelli, R., Barsh, G., Blake, J., Beffell, D., Bojunga, N.,  
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bul, C.,  
Flatcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,  
Hotmann, L., Hume, D. A., Kamlya, M., Lee, N. H., Lyons, P.,  
Machlomm, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,  
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,  
Sato, R., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,  
Toyooka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L.,  
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohzuki, S.  
and Hayashizaki, Y.

JOURNAL MEDLINE  
PUBMED  
21085651  
11217851

5 (bases 1 to 1464)  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)

JOURNAL MEDLINE  
PUBMED  
21085651  
11217851

ADACHI, J., AIZAWA, K., AKAHIRA, S., AKIMURA, T., AONO, H., ARAI, A.,  
ARAKAWA, T., BALDARELLI, R., BONO, H., BROWNSTEIN, M., BUL, C.,  
CARNINCI, P., FUKUDA, S., FUKUNISHI, Y., FURUNO, M., HANAGAKI, T.,  
HARA, A., HAYATSU, N., HILL, D., HIRAMOTO, K., HIROKA, T., HORI, F.,  
HUME, D., IMOTANI, K., ISHII, Y., ITOH, M., IZAWA, M., KASUKAWA, T.,  
KATO, H., KAWAI, J., KOJIMA, Y., KONO, H., KONDA, M., KOYA, S.,  
KURIHARA, C., MATSUYAMA, T., MIYAZAKI, A., NISHII, K., NOMURA, K.,  
NUMAZAKI, R., OHNO, M., OKAZAKI, Y., OKIDO, T., OWE, C., QUACKENBUSH, J.,  
SAITO, H., SAITO, R., SAKAI, C., SAKAI, K., SANO, H., SASAKI, D.,  
SCHIRALI, L., SHIBATA, K., SHIBATA, Y., SHINAGAWA, A., SHIRAKI, T.,  
SOGABE, Y., SUZUKI, H., TAGAMI, M., TAGAWA, A., TAKAHASHI, F.,  
TANAKA, T., TEJIMA, Y., TOYA, T., YAMAMURA, T., YAMANAKA, I.,  
YASUNISHI, A., YOSHIDA, K., YOSHINO, M., MURAMATSU, M. and  
HAYASHIZAKI, Y.

Direct Submission  
Submitted (10-JUL-2000) yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration and Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama  
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,

COMMENT Fax: 81-45-503-9216) for further details.  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGGAGAAGATCCAGACCTCTTTTCTTTTTTTTAA 3'], cDNA was prepared by using triethanolamine-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 7.5 and subtraction to Rot = 37.5. Second strand cDNA was prepared with the primer adapter of sequence [5' GAAGAGAGACTTCCTCAGATTAAATTATAATATCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstII. Cloning sites, 5' end: XhoI; 3' end: SstII. Host: SOLR.

FEATURES  
source location/Qualifiers  
1..1464  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="FANTOM.DB:2810011M06"  
/db\_xref="MGd:MGI:1901365"  
/db\_xref="taxon:10090"  
/clone="2810011M06"  
/clone\_lib="RIKEN full-length enriched mouse cdna library"  
/dev\_stage="10, 11 days embryo"  
603..1322  
/note="data source:SPTK, source key:Q9ESK4, evidence:ISS putative  
similar to P33ING2"  
/codon\_start=1  
/protein\_id="BAB28430.1"  
/db\_xref="GI:12849650"  
/db\_xref="MGd:MGI:1916510"  
/translation="MQRNVSYLRELNDKNKYQETLEIDVYERYKKEDDSNOKRLQOHLQRLINSOLGDERTIQIVQMELVENRAQMLHSQCQDPASESEASDPSKMDSSQPESRSRRPRTORSERSDLCIHMTNGIDCDOPPKERSKSARKKRSKAKORREASPVFEALDKNEPTYCLCNQOVSGEIMGDCNECPLEMFHFSCVSLTPRKWYCPCPCGDNKTMDKSTPEKTKKERRNA"  
1444..1449  
/note="putative"  
polya\_site  
1464  
/note="putative"  
BASE COUNT 370 a 333 c 483 g 278 t  
ORIGIN

Alignment Scores:  
Pred. NO.:  
Score: 5.84e-109 Length: 1464  
Best local Similarity: 1395.50 Matches: 269  
Percent Similarity: 97.86% Conservative: 6  
Query Match: 95.73% Mismatch: 5  
Indels: 2  
Gaps: 11

US-09-513-365A-1 (1-280) x AK012716 (1-1464)

Oy 1 MetLeuGly---GlnGlnGlnGlnGlnLeuTyrsSerAlaAlaLeuLeuThrGlyGlu 19  
||| |||||  
Db 478 ATGTTAGGGCGAGCGACGACACACAGCTGACCTGTGCCTGCCGCTCTCAACGGAGAG 537  
||| |||||  
Oy 20 ArgSerArgLeuLeuThrCysTyValGlnAspTryLeuGluCysValGluSerLeuPro 39  
||| |||||  
Db 538 CGGAG-CGGCGCTCTCTCTCTCTAGCTGAGGACATCACTGAGAGTGCTGGAATCTCTGCC 596  
||| |||||  
Oy 40 HisAspMetGlnArgAsnValSerValLeuArgGluLeuAspAsnLysTyrlGlnGluThr 59  
||| |||||  
Db 597 CACACACATGCGAGAGAACGTCTCGCTGCTCGCGGAGCTGGACAACAATATACAGAAAGC 656  
||| |||||  
Oy 60 LeuIysGlnIleAspAspValTyrlGluLysTyrlLysGluLysGluAspAspLeuAsnGlnLys 79  
||| |||||  
Db 657 TTAAAGCAATATGATGATGCTTAATGAAAATAATAGAGAGATGATTCACACAGAAA 716  
||| |||||

```

QY 80 LysArgLeuGlnGlnLeuGlnArgAlaLeuIleAsnSerGlnGluGlyAspGlu 99
DB 717 AAAGCGCTACAGGCACTCCACAGAGCGTTATCATACCAAGAAATGGAGATGAA 776
QY 100 LysIleGlnIleValThrGlnMetLeuGlnValGlnAsnArgAlaArgGlnMetGlu 119
DB 777 AAAATTGATGTCACCCAGATGCTCGAATTGGTGGAGAACCGAGGAGCAAAATGAG 836
QY 120 LeuHisSerGlnCysPheGlnAspProAlaGlnSerGlnArgAlaSerAspGlnAlaGly 139
DB 837 CTGATTCACAGAGTGTTCACAGATCTGCTGAAAGTAGAGCGAGCTCGACAACTCGAG 896
QY 140 MetAspSerSerGlnProGluArgSerSerArgArgProArgArgGlnArgThrSerGlu 159
DB 897 ATGATTCACAGTCAACCGAAGATCTTCTAGAAAGCTCGAAGACAGAGAGCAAGTGG 956
QY 160 SerArgAspLeuGlnCysHisMetAlaAsnGlyIleGlnAspCysAspAspGlnProProlys 179
DB 957 AGCGGTACTTATCTCAGATGACAAACGGGATTGACAGACTGTGATCAACACCGGAA 1016
QY 180 GluLysSerLysSerLysSerAlaLysLysLysLysArgSerLysLysGlnGluArgGlu 199
DB 1017 GAAAGAGATCAATCCAGTCCGCCAAGAGAGAGAGCGCTCCAGAGCGCAAGAGAGGAG 1076
QY 200 AlaSerProValGluPheAlaIleAspProAsnGluProThrTyrcysLeuGln 219
DB 1077 GCATCCCTGTCGAGTTTGCATCGATCCCAATGAGCCACCTACTGCTGTGTAACCA 1136
QY 220 ValSerTyrcysGlnMetIleGlyCysAspAsnGlnGlnCysProIleGlnTyrPheHis 239
DB 1137 GTGTCTACGCGGAGATGATAGGCTGTGACAAATGACAGTCTCCATTAATGATGTTTCC 1196
QY 240 PheSerCysValSerLeuThrTyrcysProLysGlnSerTyrcysProLysCysArg 259
DB 1197 TTTTCATGCTGTTTACCTACCTATTAACCCAGGCGAATGCTTTGCCCAAGTGTGAG 1236
QY 260 GlyAspAsnGlnLysThrMetAspLysSerThrGluLysThrLysAspArgArgSer 279
DB 1257 GGAGACAAATGAGAAACCATGAGCAAAAGTACGCAAAAGACAAAAGAGAGAGAGCG 1316
QY 280 Arg 280
DB 1317 AGG 1319

RESULT 2
LOCUS BQ277444 912 bp mRNA linear EST 07-MAY-2002
DEFINITION AGENCOURT_6738254 NIH_MGC_127 Homo sapiens CDNA clone IMAGE:5810745
ACCESSION BQ277444
VERSION BQ277444.1 GI:20487652
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 912)
NIH-MGC http://mgi.ncl.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: csapbs-remail.nih.gov
Tissue Procurement: NCI
CDNA Library Preparation: Michael Brownstein Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMN at:
http://image.llnl.gov
Plate: LICM2058 row: h column: 10
High quality sequence stop: 587.
FEATURES
Source 1. 912
Location/Qualifiers

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5810745"
/clone_lib="NIH_MGC_127"
/tissue_type="mixed (pool of 40 RNAs)"
/lab_host="DH10B (T1-phage-resistant)"
/notes="Vector: pDNR-LIB; Site: 1: SfiI (ggccattagcc); Site: 2: SfiI (ggccgctggcc); Double-stranded cDNA was prepared from a pool of 40 cell line polyA+ RNAs (bladder - 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon - 4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%, kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were used in cloning as follows:
5'-AACGATGCTTCAACGACAGAGTCCATTCAGCCGGG-3' and
5'-ATTCTAGAGCGCGGCGGCGGAGATG-dT(30)NN-3'. Full-length enriched library was constructed using the Clontech Creator SMART kit and size-selected to contain the 1-2 kb size fraction (other fractions present in NIH_MGC_126 and NIH_MGC_128). Library created in the laboratory of T. Ueda, M.D., Ph.D. (NIH, NIH). Note: this is a NIH_MGC library."
BASE COUNT 313 a 184 c 235 g 180 t
ORIGIN
Alignment Scores:
Pred. No.: 1,1e-100 Length: 912
Score: 1296.00 Matches: 261
Percent Similarity: 92.71% Conservative: 6
Best Local Similarity: 90.62% Mismatches: 11
Query Match: 87.51% Indels: 11
DB: 14 Gaps: 1
US-09-513-365a-1 (1-280) x BQ277444 (1-912)
QY 1 MetLeuGlnGlnGlnGlnGlnLeuTyrcysSerSerAlaLeuLeuThrGlyGluArg 20
DB 22 ATGTAGGCGCAGCAGCAGCAGCAACTGTACGTCGCGTGGCCCTGACCGGGAGCGG 81
QY 21 SerArgLeuLeuThrCysTyrcysValGlnAspTyrcysGlnGlnSerLeuProHis 40
DB 82 AGCGGCTGCTCACCTGCTACGTCAGAGACTCTTGTAGTGTGAGTCCGCTGCCAC 141
QY 41 AspMetGlnArgAsnValSerValLeuArgGlnLeuAspAsnLysTyrcysGlnThrLeu 60
DB 142 GACATGCGAGAGAAAGTGTCTGTCTGCGAGAGCTGGCAACAATATCAAGAAACGTTA 201
QY 61 LysGlnIleAspAspValTyrcysTyrcysGlnGlnAspAspLeuAsnGlnLysLys 80
DB 202 AAGGAANTGTATGTCTGTACGAAATATAGAAAGAGATGATTTAAACAGAGAGAA 261
QY 81 ArgLeuGlnGlnLeuGlnArgAlaLeuIleAsnSerGlnGluGlyAspGlnLys 100
DB 262 CGCTACAGCAGCTTCTCCACAGAGCACTAATTAATAGTCAAGAAATGGAGATGAAAA 321
QY 101 IleGlnIleValThrGlnMetLeuGlnLeuValGlnAsnArgAlaArgGlnMetGlnLeu 120
DB 322 ATACAGATGTTTACCAATGCTCGAATGTGTGAAAAATGGGCAAGCAAAATGAGTTA 381
QY 121 HisSerGlnCysPheGlnAspProAlaGlnSerGlnArgAlaSerAspGlnAlaGlyMet 140
DB 382 CACTCACAGTCTTCCAAAGATCCGCTGAAGTGAAGCAGCTCAGATTAAGCAAAAGTGG 441
QY 141 AspSerSerGlnProGluArgSerSerArgArgProArgArgGlnArgThrSerGlnSer 160
DB 442 GATTCACAGCCAAACAGAAAGATCTTCAAGAGACCCGCGAGCGAGCGAGCAAGTGAAGC 501
QY 161 ArgAspLeuGlnCysHisMetAlaAsnGlyIleGlnAspCysAspAspGlnProProlysGlu 180
DB 502 CGTGAATTTATGTCAATGCGCAAAATGGGATTGAAGACTGTGTGATGATCGCCACTTAAGAA 561
QY 181 LysLysSerLysSerAlaLysLysLysLysArgSerLysAlaLysGlnGluArgGluAla 200

```

```

Db      562 AAGAAATCCAGTCAGCAAGAAAGAAACGCTCCAGAGGCAAGCAAGCAAGCAAGCAAGCT 621
Qy      201 SerProValGluIleuAlaIleaspProAsnGluProThyTyrCysLeuCysAsnGlnVal 220
Db      622 TCACCGTGTAGATTGCAATAGATCCAAAGAAAGCAATACATGCTTATCCAAACCCAGT 681
Qy      221 SerTyrGluMetIleGlyCysAspAsnGluGlyCysProIleGluTyrPheHis-Ph 240
Db      682 TCTATGGGAGATGATAGATGATGACAAATGA-CAGGTCCCAATTCATGGGTTCACTTT 740
Qy      240 eSerCysValSerLeuThyTyrLys-ProLysGlyLysTyrPyr-CysProLysCysArg 259
Db      741 TTCATGTGTTCACTTACCTATTAACCCAAAGGGAATGCTATTGTCACCAAGTGCAGG 800
Qy      260 -GlyAspAsn-----GluTyrThrMetAspLysSerThrGluLysThr 273
Db      801 GGAGATTAATGAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 860
Qy      273 rLysLysAspArgArg 278
Db      861 TAGAAAAAGATCCAAAG 876

RESULT 3
Bg211544 919 bp mRNA linear EST 21-APR-2001
LOCUS RST31111 Atherys RAGE Library Homo sapiens cDNA, mRNA sequence.
DEFINITION Bg211544
ACCESSION Bg211544.1 GI:13733231
VERSION EST.
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
            Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            1 (bases 1 to 919)
REFERENCE
AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
            Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.,
            Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith
            ,E., Veloso,N., Kilka,A., Hess,J., Cothren,K., Lo,K., Offenbacher
            ,J., Danzig,J. and Ducar,M.
            Creation of genome-wide protein expression libraries using random
            activation of gene expression
            Nat. Biotechnol. 19 (5), 440-445 (2001)
JOURNAL MEDLINE 21227151
COMMENT Contact: Scott J. Cain
            Atherys, Inc.
            3201 Carnegie Ave, Cleveland, OH 44115, USA
            Tel: 216 431 9800
            Fax: 216 361 9596
            Email: scain@atherys.com
            High quality sequence stop: 439.
            Location/Qualifiers
FEATURES
source 1. 919
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_lib="Atherys RAGE Library"
            /cell_line="HT1080"
            /note="See 'Creation of Genome-wide Protein Expression
            Libraries using Random Activation of Gene Expression',
            Nature Biotechnology, in press. Note that even though the
            cell type indicated is HT1080, since a random activation
            method was used, these sequence tags are not necessarily
            expressed in HT1080 under normal circumstances."
BASE COUNT 332 a 151 c 216 g 216 t 4 others
ORIGIN
Alignment Scores:
Pred. No.: 1.04e-83 Length: 919
Score: 1096.50 Matches: 214
Percent Similarity: 96.00% Conservative: 2
Best Local Similarity: 95.11% Mismatches: 7
Query Match: 74.04% Indels: 4
DB: 12 Gaps: 1

```

```

US-09-513-365a-1 (1-280) x Bg211544 (1-919)
Qy      58 GluThrLeuLysGluIleaspAspValTyrGluLysTyrLysGluAspAspLeuAsn 77
Db      4 GAAACGTTAAAGCAAAATTGATGATGCTTACGAAAAATATAGAAAGAGATGATTTAAAC 63
Qy      78 GluLysLysArgLeuGluGluIleuLeuGluAlaLeuIleAsnSerGlnGluLeuGly 97
Db      64 CAGAGAACGCTTACAGCAGCTTCTCCAGAGAGCACTTAATTAATGTCAGAAATGGGA 123
Qy      98 AspGluLysIleGluIleValThrGluMetLeuGluLeuValGluAsnArgAlaArgGln 117
Db      124 GATGAATAAATACAGATTGTTACACAATGCTCGAATTTGGTGGAAATCGGCAAGCA 183
Qy      118 MetGluLeuHisSerGlnCysPheGlnAspProAlaGluSerLysLysAspAspLys 137
Db      184 ATGAGATTACACTACACAGTGTTCACAGATCCCTGCTGAAAGTAAAGAGAGCTCAGATAAA 243
Qy      138 AlaLysMetAspSerSerGlnProGluArgSerSerArgProArgArgGlnArgThr 157
Db      244 GCAAAGTGTGATTCAGCCACCAAGAAAGATCTTCAGAGAGACCCGACAGCGAGCC 303
Qy      158 SerGluSerArgAspLeuCysHisMetAlaAsnGlyIleGluAspCysAspAspGlnPro 177
Db      304 AGTGAAGCCGCGTATTTATGTCACATGCGCAATGGGATGAGACGTGATGATCAGCCA 363
Qy      178 ProLysGluLysSerLysSerAlaLysLysLysLysArgSerLysAlaLysGlnGlu 197
Db      364 CTTAAAGAAAGAAATCCCAAGTCCAGCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 423
Qy      198 ArgGluAlaSerProValGluPheAlaIleaspProAsnGluProThyTyrCysLeuCys 217
Db      424 AGGGAAGCTTCACCTGTGTGATTTGCATATGATCCTAATGACCTATGCTATATGC 483
Qy      218 AsnGlnValSerTyrGluMetIleGlyCysAspAsnGluGlnCysProIleGluTyrP 237
Db      484 AACCAAGTCTTATGGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 543
Qy      238 PheHisPheSerCysValSerLeuThrTyrLysProLysGlyLysTyrPyrCysProLys 257
Db      544 TTTCACCTTTCACTGTTCTTCACTTACCTATTAACCAAGGGAAGGGAATGGATCCCANAG 603
Qy      258 CysArgGluAspAsnGluLysThrMetAspLysSerThrGluLysThrLysLys----- 275
Db      604 TCCAGGCGAGCAATATGAGAAACAAA-GACAAAGC-ACGAAAGAAAGCAAAAGATGAGAA 661
Qy      276 AspArgArgSerArg 280
Db      662 GATCGAAGTAAAG 676

RESULT 4
Bg796780 793 bp mRNA linear EST 20-SEP-2000
LOCUS BE796780
DEFINITION 60158757F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3941655 5',
ACCESSION BE796780
VERSION BE796780.1 GI:10218080
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
            Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            1 (bases 1 to 793)
REFERENCE
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabs-r@mail.nih.gov
            Tissue Procurement: DCTD/DRP
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
            DNA Sequencing by: Incyte Genomics, Inc.

```



```

Db      184 AGCCGGCTGCTACCTGCTACGTGACAGGACTACCTTGAGTGGCTGGAGTCCGTCACAC 243
QY      41 ASPMeGlnArgAsnValSerValLeuArgGluLeuAspAsnIlySerGlnGluThrLeu 60
Db      244 GACATCGACAGCAACGTGTGTGCTGCGGAGACTGGACACAAATATCAAGAAACGTTA 303
QY      61 LysGluIleAspAspValTyrgIuLysTyrgLysGluAspAspLeuAsnGlnIlyLys 80
Db      304 AAGCAATGTATGATGCTACGAAATAATATAGAAAGAAAGATGATTTAAACGAGAGAAA 363
QY      81 ArgLeuGlnGlnLeuLeuArgGlnAlaLeuIleAsnSerGlnGlnLeuGlnGlnGlnLys 100
Db      364 CGTTCACAGCGCTTCCAGAGACACTAATTAATGCTAAGAAATGGAGATGAAAAA 423
QY      101 IleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAlaArgGlnMetGluLeu 120
Db      424 ATACAGATTGTTACCAAAATGCTCGAATGCTGTAAGAAATGGCCAGACAAATGAGATT 483
QY      121 HisSerGlnCysPheGlnAspProAlaGluSerGlnArgAlaSerAspIlyAlaLysMet 140
Db      484 CACTCAGAGTGTCCAGAGATCCTGCTGAAGTGAACGAGCCTCAGATAAGCAAGATG 543
QY      141 AspSerSerGlnProGlnArgSerArgArgProArgArgGlnArgGlnArgGlnSerGln 160
Db      544 GATTCACACCAACCAAGAAATCTTCAAGAGACCCGCGAGCGAGCGACAGTGAAGC 603
QY      161 ArgAspLeuGlnHisMetAlaAsnGlnIleGluAspCysAspAspGlnProProLysGlu 180
Db      604 CGTGATTTATGTCACATGGCAATGGATGGAAGATGAGAGCTGATGATGACCACTAAAGAA 663
QY      181 LysLysSerLysSerLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 200
Db      664 AAGAAATCCCAAGTCACCAAGAAAGAAAGCGCTCCAGAGC -AAGCAGGAAAGGAGCT 722
QY      201 SerProValGluPheAlaIleAspProAsnGlnProThrTyrcLysCysAsnGln 219
Db      723 TCACCTGTGAGTTGTC -ATAGATCCTAATGAACCTACATCTGCTTATGCAACAAG 778

RESULT 6
LOCUS   BM982877 735 bp mRNA linear EST 21-MAR-2002
DEFINITION
UI-CF-EN1-acs-d-05-0-UI-s1 UI-CF-EN1 Homo sapiens cDNA clone
UI-CF-EN1-acs-d-05-0-UI 3', mRNA sequence.
ACCESSION
BM982877
VERSION
BM982877.1 GI:19606826
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 735)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL
MEDLINE
97044477
COMMENT
Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 FORWARD
POLYA-yes.
FEATURES
source
1..735
Location/Qualifiers

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-CF-EN1-acs-d-05-0-UI"
/clone_1ib="UI-CF-EN1"
/tissue_type="Primary Lung Cystic Fibrosis Epithelial
Cells"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/Note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-CF-EN1 is a normalized cDNA library containing the
following tissue(s): Primary Lung Cystic Fibrosis
Epithelial Cells. The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was ligated to an EcoR I adaptor, digested with Not
I, and cloned directionally into pT73-Pac vector. The
oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is CAGCTCAGCT.
TAG_Lib-UI-CF-EN1
TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS
6hr to LPS_24hr
TAG_SEQ=CTGCTCAGCT"
BASE COUNT 180 a 161 c 114 g 280 t
ORIGIN
Alignment Scores:
Pred. No.: 1,25e-77 Length: 735
Score: 1024.00 Matches: 190
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 69.14% Indels: 0
DB: 14 Gaps: 0
US-09-513-365A-1 (1-280) x BM982877 (1-735)
QY 91 IleAsnSerGlnGlnLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 110
Db 735 ATTAATATGTCACAAATTTGGAGATGAGAAAAATACATGTTTTCACAAATGCTCGAATG 676
QY 111 ValGluAsnArgAlaArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGlu 130
Db 675 GTGGAATAATCGGCGCAAGCAAAATGAGATTACACTCAGAGTGTCCAAATCCTCGTCAA 616
QY 131 SerGlnArgAlaSerAspIlyAlaLysMetAspSerSerGlnProGlnArgSerArg 150
Db 615 AGTGAACGAGCCCTCAGATTAAGCAAGATGATTCAGCCACAGAAAGATCTTCAAGA 556
QY 151 ArgProArgGlnArgThrSerGlnSerArgAspLeuGlnHisMetAlaAsnGlnIly 170
Db 555 AGACCCCGCAGAGCGGACAGTGAAGCCGATTTATGTCACATGGCAATGGGATT 496
QY 171 ArgAspCysAspAspGlnProProLysGlnLysLysSerLysSerAlaLysLysLys 190
Db 495 GAAGACTGTGATGATCAGCACCTTAAGAAAGAAATCAAGTCACACAAAGAAAGAA 436
QY 191 ArgSerLysAlaLysGlnGlnArgGluAlaSerProValGluPheAlaIleAspProAsn 210
Db 435 CGCTCCAGGCGCAAGAGAGGAGGAGGCTTCCCTGTTGAGATTGCAATAGATCCTAAT 376
QY 211 GluProThrTyrcLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 230
Db 375 GAACCTACATACGCTTATGACACCAAGTGTCTTATGGGAGATGATAGATGTCAAT 316
QY 231 GluGlnCysProIleGlnIlyPheHisPheSerCysValSerLeuThrTyrgLysProLys 250
Db 315 GAACAGTGTGCAATGATGATTCACTTTTCACTTGTGTTCACTTACTATTAACCAAG 256
QY 251 GlyLysTrpTyrcLysProLysCysArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 270

```

Db 255 GGGAAATGGTATTGCCCAAGTCAGGAGGATATGAGAAACAATGACAAAGACTACT 196  
 QY 271 GGUUThrlslyspargargserarg 280  
 Db 195 GAAAGACAAAGATGAGATGAGATGAGG 166  
 RESULT 7  
 BG184056/c 798 bp mRNA linear EST 21-Apr-2001  
 LOCUS BG184056  
 DEFINITION RST2972 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BG184056  
 VERSION BG184056.1 GI:13705743  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 REFERENCE 1 (bases 1 to 798)  
 AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,  
 Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.,  
 Lerner,L., Costanzo,D., McElligott,K., Boozar,S., Mays,R., Smith,  
 E., Veloso,N., Kijka,A., Hess,J., Cothren,K., Lo,K., Offenbacher,  
 J., Danzig,J. and Ducar,M.  
 TITLE Creation of genome-wide protein expression libraries using random  
 activation of gene expression  
 JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)  
 MEDLINE 21227151  
 COMMENT Contact: Scott J. Cain  
 Athersys, Inc.  
 3201 Carnegie Ave, Cleveland, OH 44115, USA  
 Tel: 216 431 9900  
 Fax: 216 361 9596  
 Email: scain@athersys.com  
 High quality sequence stop: 531.  
 Location/Qualifiers  
 1..798  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="Athersys RAGE Library"  
 /cell\_line="HT1080"  
 /note="See 'Creation of Genome-wide Protein Expression  
 Libraries using Random Activation of Gene Expression',  
 Nature Biotechnology, in press. Note that even though the  
 cell type indicated is HT1080, since a random activation  
 method was used, these sequence tags are not necessarily  
 expressed in HT1080 under normal circumstances."  
 BASE COUNT 197 a 178 c 130 g 293 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 4.15e-76 Length: 798  
 Score: 1006.50 Matches: 209  
 Percent Similarity: 95.43% Conservative: 0  
 Best Local Similarity: 95.43% Mismatches: 10  
 Query Match: 67.96% Indels: 6  
 Gaps: 0  
 US-09-513-365a-1 (1-280) x BG184056 (1-798)  
 QY 62 GGUUThrlslyspargargserarg 280  
 Db 798 GAAATGGTATTGCCCAAGTCAGGAGGATATGAGAAACAATGACAAAGACTACT 741  
 QY 82 GGUUThrlslyspargargserarg 280  
 Db 740 CTCAGCAGCTTCTCCAGAGACCTATTAT-AGTCAAGCAATGGGAGATGAAAAA- 683  
 QY 102 GGUUThrlslyspargargserarg 280  
 Db 682 CAAATGTTTCA-CAATGCTCGAATGCTGGA-ATGCGGCAAGACAAATGAGTTTCA 625  
 QY 122 SerGlnCysPheGlnAspProAlaGluSerGluArgAlaSerAlaLysMetAsp 141

Db 624 TCACAGTGTTCACAGATCCTGCTGAAGTGAACGAGCCCTCAGATAAAGCAAGATGAT 565  
 QY 142 SerSerGlnProGluArgSerArgArgProArgGlnArgGlnArgGlnArgGlnArg 161  
 Db 564 TCACAGCCACCAAGAAATGATCTCAAGAAAGCCCGGAGGAGACCACTGAAAGCCGT 505  
 QY 162 AspleucyshiMetalaansgylileglaaspCysaspaspGlnProPolysGluLys 181  
 Db 504 GATTATGTCACATGCGCAATGGGATGGAAGACTGATGATGATGATGATGATGATGAT 445  
 QY 182 LysSerLysSerAlaLysLysLysLysArgSerLysAlaLysGlnGlnArgGlnAlaSer 201  
 Db 444 AAATCCAAAGTCACCAAGAAAGAAAGAAAGCCCTCCAGAGCCCAAGCAAGAAAGGAGCTTCA 385  
 QY 202 ProValGlnPheAlaIleAspProAsnGluProThrTrpTrpLysLysGlnGlnValSer 221  
 Db 384 CCGTGTGATTTGCAATAGATCCCAAGAACTCAATACGCTTATCAACCAAGCTCTC 325  
 QY 222 TyTGlyGlnMetilegLYCysaspaspGlnGlnCysProilegIutRppheAspSer 241  
 Db 324 TATGGGAGATGATAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 265  
 QY 242 CysValSerLeuThrTrpLysProLysGlyLysTrpTrpCysProLysCysArgGlyAsp 261  
 Db 264 TGTGTTCACTTACTTAAACCAAGGAGAAAGTGTATGCCCAAGTGCAGGGAGAT 205  
 QY 262 AsnGluLysThrMetAlaLysSerThrGluLysThrLysLysAspArgArgSerArg 280  
 Db 204 AATGAGAAACAAATGACAAAGTACTGAAAGCAAAAGATGATGATGATGATGATGATGAT 148  
 RESULT 8  
 BM719460 664 bp mRNA linear EST 01-MAR-2002  
 LOCUS BM719460  
 DEFINITION UT-E-EJ1-ajf-g-22-0-UT-1 UT-E-EJ1 Homo sapiens cDNA clone  
 UT-E-EJ1-ajf-g-22-0-UT 5', mRNA sequence.  
 ACCESSION BM719460  
 VERSION BM719460.1 GI:19038113  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 REFERENCE 1 (bases 1 to 664)  
 AUTHORS Bonaldo,M.F., Denon,G. and Soares,M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 COMMENT Contact: Soares, MB  
 Program for Rat Gene Discovery and Mapping  
 University of Iowa  
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: msoares@blue.weeg.uiowa.edu  
 Tissue Procurement: Dr. Gregg Hageman  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Genome Distribution: Researchers may obtain clones from Research  
 Genetics (www.resgen.com).  
 Seq primer: M13 Reverse.  
 FEATURES  
 source  
 1..664  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="UT-E-EJ1-ajf-g-22-0-UT"  
 /clone\_lib="UT-E-EJ1"  
 /tissue\_type="fetal eyes, lens, eye anterior segment,  
 optic nerve, retina, Retina foveal and Macular, RPE and  
 Choroid"  
 /dev\_stage="fetal and adult"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"

/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site\_1: EcoR I; Site\_2: Not I; UI-E-E11 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are: fetal eyes, AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment, AATGCCGAT; optic nerve, CCATTAGTG; retina, CCGCG; Retina foveal and macular, GTCC; RPE and Choroid, ACCGA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

BASE COUNT 241 a 111 c 149 g 162 t 1 others

## ALIGNMENT SCORES:

Pred. No.:	Length:	664
Score:	Matches:	177
Percent Similarity:	Conservative:	0
Best Local Similarity:	Mismatches:	0
Query Match:	Indels:	0
DB:	Gaps:	0

US-09-513-365a-1 (1-280) x BW19460 (1-664)

QY 104 ValThhGlmleuclueuValGluAsnAlaAlaArgGlmleuHISserGln 123  
 DB 1 GTTACCAATGCTCCAAATGTGTGGAATCGGGCAAGAACAAATGAGTACACTACAG 60  
 QY 124 CysPhgGlnAspProAlaGlnSerGluArgAlaSerAspLysMetAspSer 143  
 DB 61 TGTTCAGAGATCTCTGAAAGTGAACGAGCTCGATTAACAAAGATGATTCACG 120  
 QY 144 GlnProGlnArgSerArgProArgGlnArgGlnSerGlnSerArgAspLeu 163  
 DB 121 CAACGAAAGATCTTCAAGAAAGACCCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180  
 QY 164 CysHisMetAlaAsnGlyIleGluAspCysAspAspGlnProPolysGluLysSer 183  
 DB 181 TGTCAATGCAATGGATGGAAGCTGATGATGATGATGATGATGATGATGATGATG 240  
 QY 184 LysSerAlaLysLysLysLysArgSerLysAlaLysGlnGlnArgGlnAlaSerPro 203  
 DB 241 AAGTCAGCAAGAAAGAAAGAAAGCTCCAGAGCCAGAGAGAGAGAGAGAGAGAG 300  
 QY 204 GluPheAlaIleAspProAsnGlnProThrTyrcysLeuGlnValSerTyrgly 223  
 DB 301 GAGTTGCAATAGATCTTATGAACTACATCTGTTATGCAACCAAGTGTCTTATGG 360  
 QY 224 GluMetIleGlyAspAsnGlnGlnGlnProIleGluTrpPheHisPheSerCysVal 243  
 DB 361 GAGATATAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420  
 QY 244 SerLeuThrTyrlsProLysGlyLysTrpTyrcysProLysCysArgGlyAspAsnGlu 263  
 DB 421 TCTCTACCTATTAACCAAGGGAATGTATGTCCCAAGAGTCCAGGAGGAGTAATGAG 480  
 QY 264 LysThrMetAspLysSerThrGluLysThrLysLysAspArgArgSerArg 280  
 DB 481 AAAACAATGACAAAGTACTGAAAGACAAAAGATGAAAGATGAAAGATGAAAGATG 531

RESULT 9  
 BI460319 708 bp mRNA linear EST 21-AUG-2001  
 LOCUS 603201967F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:5267729 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BI460319  
 VERSION BI460319.1 GI:15250975

KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 708)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki Toshiyuki and Piero Carninci (RIKEN)  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Invitrogen, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>  
 Plate: L14M1675 row: f column: 18  
 High quality sequence stop: 700.

## FEATURES

source  
 Location/Qualifiers  
 1..708  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5267729"  
 /clone\_1ib="NIH\_MGC\_97"  
 /lab\_host="DH10B"

/note="Organ: testis; Vector: pBluescript (modified pBluescript KS+); Site\_1: BamHI; Site\_2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTVA-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH-MGC Library."

BASE COUNT 209 a 163 c 213 g 122 t 1 others

US-09-513-365a-1 (1-280) x BI460319 (1-708)

QY 1 MetLeuGlyGlnGlnGlnGlnLeuTyrsSerAlaAlaLeuThrGlyGluArg 20  
 DB 139 ATGTTAGGGAGAGAGAGAGAGCAAGTACTGCGGCGGCTCTGAGCGGAGCGG 198  
 QY 21 SerArgLeuThrCysTyrlsAlaGlnAspTyrlsGlnGlyValGlnSerLeuProHis 40  
 DB 199 AGCGGCTGCTCCCTGCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 258  
 QY 41 AspMetGlnArgAsnAlaSerValLeuArgGlnLeuAspAsnLysTyrlsGlnGluThrLeu 60  
 DB 259 GACATCAG 318  
 QY 61 LysGluIleAspAspValTyrlsGluTyrlsLysGlnAspAspLeuAsnGlnLysLys 80  
 DB 319 AAGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 378  
 QY 81 ArgLeuGlnGlnLeuLeuGlnArgAlaLeuIleAsnSerGlnGlnGlnGlnGlnGln 100  
 DB 379 CGTCTACAG 438  
 QY 101 IleGlnIleValThrGlnMetLeuGlnLeuValGluAsnArgAlaArgGlnMetGluLeu 120  
 DB 439 ATACAGATTGTACACAAATGCTGGAATTGTGGAATGCGGCAAGACAAATGAGATT 498



QY 121 HisserglnCysPheGlnAspProAlaGluSerGluArgAlaSerAspLysAlaLys-Me 140  
 DB 499 CACACACAGCTTTCCAGATCTCTGCTAAGAGTACAGCCTCGATTAAGCAAGATT 558  
 QY 140 TASPserSerGlnProGluArgSerSerArgArgProArgGlnArgThrSerGlu 160  
 DB 559 GGATTCGAGCCACAGAAAGATCTTCAAGAAAGACCCGCGAGCGGACAGTGAAG 618  
 QY 160 rArAspLeuGlnHisMetAlaAsnGlyTlGlnAspCysAspGlnProProLysG 180  
 DB 619 CCGGATTTATGATCATCATGCAATGCGAATGGATGAAGACTGATGATNCAGCAGCTAAAGA 678  
 QY 180 ULVSLSerLysSerAlaLysLysLys 189  
 DB 679 AAGAAATCCAGATCAGCAAGAAAG 706  
 RESULT 10  
 BM682789/c 675 bp mRNA linear EST 27-FEB-2002  
 LOCUS UT-E-E11-ajf-g-22-0-UI s1 UT-E-E11 Homo sapiens cDNA clone  
 DEFINITION UT-E-E11-ajf-g-22-0-UI 3', mRNA sequence.  
 ACCESSION BM682789  
 VERSION BM682789.1 GI:18992685  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 675)  
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 9704447  
 COMMENT Contact: Soares, MB  
 Program for Rat Gene Discovery and Mapping  
 University of Iowa  
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9365  
 Email: msoares@blue.weeg.uiowa.edu  
 Tissue Procurement: Dr. Gregg Hageman  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research  
 Genetics (www.resgen.com).  
 Seq primer: M13 Forward  
 POLY-A-yes.  
 FEATURES  
 SOURCE location/Qualifiers  
 1. 675  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="UI-E-E11-ajf-g-22-0-UI"  
 /clone\_lib="UI-E-E11"  
 /tissue-type="fetal eyes, lens, eye anterior segment,  
 optic nerve, retina, Retina Foveal and Macular, RPE and  
 Choroid"  
 /dev-stage="fetal and adult"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a  
 modified polylinker; Site\_1: EcoR I; Site\_2: Not I;  
 UI-E-E11 is a subtracted cDNA library constructed  
 according to Bonaldo, Lennon and Soares, Genome Research,  
 6:791-806, 1996. First strand cDNA synthesis was primed  
 with an oligo-dT primer containing a Not I site. Double  
 stranded cDNA was ligated to an EcoR I adaptor, digested  
 with Not I, and cloned directionally into pT73-Pac  
 vector. The oligonucleotide used to prime the synthesis of  
 first strand cDNA contains a library tag sequence that is  
 located between the Not I site and the (dT)18 tail. The  
 sequence tags for this library are: fetal eyes, AGAATCAAGA

; lens, CGATTAGCGA; eye anterior segment, AATGCCGAT;  
 optic nerve, CCATTAGTG; retina, CCGG; Retina Foveal and  
 Macular, GTCC; RPE and Choroid, ACCTA. This library was  
 created for the program, Gene Discovery in the Visual  
 System, supported by National Eye Institute (NEI).  
 TAG\_Lib=UI-E-E11  
 TAG\_Tissue=Foveal and Macular Retina  
 TAG\_SEQ=GTCC"  
 BASE COUNT 164 a 148 c 110 g 253 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 8,12e-69 Length: 675  
 Score: 920.00 Matches: 168  
 Percent Similarity: 99.41% Conservative: 1  
 Best Local Similarity: 98.82% Mismatches: 1  
 Query Match: 62.12% Indels: 0  
 DB: 14 Gaps: 0  
 US-09-513-365a-1 (1-280) x BM682789 (1-675)  
 QY 111 ValGluAsnArgAlaArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGlu 130  
 DB 675 GTGGAAATCCGGGCAAGCAAAATGGAGTTACATCACCCTGTTCCAGATCTGCTGAA 616  
 QY 131 SerGluArgAlaSerAspLysAlaLysMetAspSerSerGlnProGluArgSerArg 150  
 DB 615 AGTGAACAGAGCTCTGATTAAGCAAGATGGATTCACCAACCAAGAAAGATCTTCAAGA 556  
 QY 151 ArgProArgArgGlnArgThrSerGluSerArgAspLeuGlnHisMetAlaAsnGlyTle 170  
 DB 555 AGACCCCGCAGCGAGCGAGCCAGTAAAGCCGTGATTTATGTCATATGGCAATGGATT 496  
 QY 171 GluAspCysAspAspGlnProProLysGluLysSerLysSerAlaLysLysLys 190  
 DB 495 GAAGACTGTGATGATCAGCCACCTAAAGAAAGAAATCCAGTCAGCAAGAAAGAA 436  
 QY 191 ArgSerLysAlaLysGlnGluArgGluAlaSerProValGluPheAlaLysAspProAsn 210  
 DB 435 CGCTCCAGAGGCCCAAGCAAGAAAGGAAACCTTCCTGTTGATGATGATGATGAT 376  
 QY 211 GluProThrTyrCysLeuGlnCysAsnGlnAlaSerTyrGlyGluMetLleGlyCysAspAsn 230  
 DB 375 GAACCTACATATCTGTTATGCAACCAATGCTTATGGAGATGATGATGATGAT 316  
 QY 231 GluGlnCysProLleGluThrPheHisPheSerCysValSerLeuThrTyrLysProLys 250  
 DB 315 GAACAGTCTCCAAATGGAATGGTTTCATCTGATGTTTCACTTACCTATAAACCAAG 256  
 QY 251 GlyLysTrpTyrCysProLysCysArgGlyLysAspAsnGlyLysThrMetAspLysSerThr 270  
 DB 255 GGGAAATGATATGCGCCAAAGTGCAGGGGAGATTAATGAGAAACATGACAAAGTACT 196  
 QY 271 GluLysThrLysLysAspArgSerArg 280  
 DB 195 GAAAGACAAAAAAGGATAGAAAGATCGAGG 166  
 RESULT 11  
 BI091047/c 891 bp mRNA linear EST 20-JUN-2001  
 LOCUS 602855235F1 NIH\_MGC\_10 Homo sapiens cDNA clone IMAGE:4996593 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BI091047  
 VERSION BI091047.1 GI:14509377  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 891)  
 AUTHORS NIH-MGC http://mhc.nci.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: AFCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: Incyte Genomics, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.lnl.gov  
plate: LLM11022 row: e column: 10  
High quality sequence stop: 722.

FEATURES  
source  
Location/Qualifiers  
1..891  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4996593"  
/clone\_1lib="NH\_MGC\_10"  
/cell\_line="MGC36"  
/lab\_host="DH10B"  
/note="Organ: cervix; Vector: pCMV-SPORT6; Site: 1: NotI; Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."

BASE COUNT 237 a 199 c 149 g 306 t

ORIGIN

Alignment Scores:  
Pred. No.: 3,58e-67 Length: 891  
Score: 902.00 Matches: 192  
Percent Similarity: 85.11% Conservative: 8  
Best Local Similarity: 81.70% Mismatches: 25  
Query Match: 60.90% Indels: 11  
DB: 13 Gaps: 2

US-09-513-365a-1 (1-280) x BI091047 (1-891)

QY 56 TrrGlnGlnThrLeuLysGlnLleAspAspValTrrGlnLysTrrLysLys---GluAs 74  
DB 829 TTTGAGGAAAGCTGTAAATGAAATTTCTACGTCATTAATTAATGATAGACA 770  
QY 74 PAsp-LeuAsnGlnLysArg-----LeuGlnLleuLeuGlnArgAla-Lleu 90  
DB 769 TGAATGTTAAACCCAGAGAGAGAGCGTTCTACAGCCAGTCTTCTCCAGAGACCTTA 710  
QY 91 IleAsnSerGlnLleuGlnLysArgGlnLysLleGlnLleVal-ThrGlnMetLeuGln 110  
DB 709 ATTAATAGTCACAAATTTGGAGTATGATTAATACAGATTGACACTATGCTCGAAT 650  
QY 110 UValGlnAsnArgAlaArgGlnMetGlnLleuHisSerGlnCysPheGlnAspPro-AlaG 130  
DB 649 GGTGGTTAAACGCA-AGACATATGAGATTACACTCACAGTGTTCACAGATCTGTCTG 591  
QY 130 lueSerGlnArgAlaSerAspLysAlaLysMet-AspSerSerGlnProGlnArgSerSer 149  
DB 590 AAAGTACAGAGCTCAGATTAAGCAAAAGTGGATTCAGCCAGCAAGAAATCTTCA 531  
QY 150 ArgArgProArgArgGlnArgLysThrSerGlnSerArgAspLeuGlnLysHisMetAla 169  
DB 530 AGAAGACCCGCGAGCGAGCGACAGTGAAGCCGCGATTTATGTCACATGCGCAATGGG 471  
QY 170 lIleGlnAspCysAspAspGlnProProLysGlnLysSerLysSerAlaLysLysLys 189  
DB 470 ATTGAAGACTGTATATATCATGCCACTAAAGAAAGAAACCAAGTCCAGCAAGAAAG 411  
QY 190 LysArgSerLysAlaLysGlnLysArgGlnAlaLysSerProValGlnPheAlaLleAspPro 209  
DB 410 AAACGCTCCAGGCGCAAGCGAAGGAAGAGCTTCACCTTTAGAGTTTGCATAATAGATCT 351  
QY 210 AsnGlnProThrTrrCysLeuGlnValSerTrrGlnLleuMetLleGlnCysAsp 229  
DB 350 AATGAACCTACATAGCTTATGCAACCAAGTGTCTTATGGGAGATGATGATGAGAC 291  
QY 230 AsnGlnGlnCysProLleGlnLysPheHisPheSerCysValSerLeuThrTrrLysPro 249

DB 290 ACTGAACACTGTCACATTAAGTGTTCACCTTTCATGCTGTTCACTTACCTAATAACCA 231  
QY 250 LysGlnLysTrrPrrCysProLysCysArgGlnLysAspAsnGlnLysThrLleLysSer 269  
DB 230 AAGGGAAATGATTTGCCCCAAGTCCAGGGGAGATTAATGACAAACAAATGACCAAAAG 171  
QY 270 ThrGlnLysThrLysLysAspArgSerArg 280  
DB 170 ACTGAAGACCAAAAGATGATGATGAGAG 138

RESULT 12  
BF114684  
LOCUS  
DEFINITION  
7566c08.x1 Soares NSF\_F8\_9W\_OT\_PA\_P\_S1 Homo sapiens cDNA clone  
IMAGE:3391406.3 similar to TR-095698 095698 INCLIP.; mRNA  
sequence.  
ACCESSION  
BF114684  
VERSION  
BF114684.1 GI:10984171  
KEYWORDS  
SOURCE  
ORGANISM  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 541)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
JOURNAL  
Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
This clone is available royalty-free through LNLN; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
Seq primer: -40UP from Gibco  
High quality sequence stop: 459.

FEATURES  
source  
Location/Qualifiers  
1..541  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3391406"  
/clone\_1lib="Soares\_NSF\_F8\_9W\_OT\_PA\_P\_S1"  
/lab\_host="DH10B"  
/note="Organ: pooled; Vector: p773p-Pac (Pharmacia) with  
a modified polylinker; Site: 1: Not I; Site: 2: Eco RI;  
Equal amounts of plasmid DNA from five normalized  
libraries were mixed, and ss circles were made in vitro.  
Following HAP purification, this DNA was used as tracer in  
a subtractive hybridization reaction. The driver was  
PCR-amplified cDNAs from pools of 5,000 clones made from  
the same 5 libraries. The pools consisted of the following  
libraries and cloneIDs: Soares NBHSF pool 1:  
309384-310919, 323208-325895 Soares NB2HP pool 1:  
145032-147335, 147720-148103, 148872-149255, 15002 -  
150407, 151176-152327 Soares NB2HF-9W pool 1:  
758280-760583, 772104-774407 Soares NBHPA pool 1:  
304776-306311, 320136-322823, 326280-326663 Soares NBHOF  
pool 1: 723720-726407, 739080-740999 Subtraction by Bento  
Soares and M. Fatima Bonaldo."

BASE COUNT 178 a 113 c 145 g 105 t

ORIGIN

Alignment Scores:  
Pred. No.: 5.98e-67 Length: 541  
Score: 897.00 Matches: 175  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.43% Mismatches: 0  
Query Match: 60.57% Indels: 0  
DB: 12 Gaps: 0

US-09-513-365a-1 (1-280) x BF114684 (1-541)

QY 1 MetLeuGlnGlnGlnGlnGlnLysSerSerAlaLleuThrGlnGlnArg 20  
|||||

Db		14	ATGTTAGGGCAGCAGACAGCAACTGACTGTCTGCCTGCCATTCCTGACCGGGGAGCGG	73
Oy		21	SerArgLeuLeuThrCysTyrValGlnAspTyrLeuGlyCysValGluSerLeuProHis	40
Db		74	AGCCGGCGTGCTCACCTGCTACGTACAGACTACCTTGAGTGTGCGAGATCGCTGCCACC	133
Oy		41	AspMetGlnArgAsnValSerValLeuArgGluLeuAspAsnLysTyrGlnGluThrLeu	60
Db		134	GACATGCAGAGAAGAAAGTGTCTGTCTGTCTGGAAGGTGGCAACAATAATCAAGAAACGTTA	193
Oy		61	LysGluIleLeuAspAspValTyrGluLysTyrLysGluAspAspLeuAsnGlnLysLys	80
Db		194	AAGGAANTGTGATGTCTACGAAAATATTAAGAAAGATGATTTTAACACAGAACAA	253
Oy		81	ArgLeuGlnGlnLeuLeuGlnArgAlaLeuIleAsnSerGlnGlnLeuGlyAspGluLys	100
Db		254	CGTCTACAGCAGCTTCTCCAGAGAGCACTAATTATATAGTCAAGAATTGGAGATGAAAAA	313
Oy		101	IleGlnIleValIleThrGlnMetLeuGlnLeuValGluAsnArgAlaIleGlnMetGluLeu	120
Db		314	ATPACAGATGTGTACACAATCTCTGAAATGGTGGAAAAATCGGGCAAGCAATGGAGTTA	373
Oy		121	HisSerGlnCysPheGlnAspProAlaGluSerGluArgAlaSerAspLysAlaLysMet	140
Db		374	CACTCACAGCTTTTCCAAAGATCTCGTGAAGTAGACAGACCTCAGATTAAGCAAAGTG	433
Oy		141	AspSerSerGlnProGluArgSerSerArgArgProArgArgGlnArgThrSerGluSer	160
Db		434	GATTCACAGCCAAACAGAAAGATCTTCAAGAAGACCCCGCAGCGAGCCAGCATGAAAGC	493
Oy		161	ArgAspLeuGlyHisMetAlaAsnGlyIleGluAspCysAspAspGln	176
Db		494	CGTATATTATGTACATGCGAAATGGATGAAAGACTGTGATGATCAG	541

RESULT\_13

LOCUS	AV647917	679 bp	mRNA	linear	EST_15-JAN-2002
DEFINITION	AV647917	GLC Homo sapiens cdna clone GLCBD03 3', mRNA sequence.			
ACCESSION	AV647917				
VERSION	AV647917.1	GI:9868931			
KEYWORDS	EST.				
SOURCE	Homo sapiens				
ORGANISM	human.				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
TITLE	1 (bases 1 to 679)				
JOURNAL	Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,				
MEDLINE	Xiao,H., Ou,Y., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,				
COMMENT	Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,				
	Hu,G., Gu,J., Chen,Z., and Han,Z.				
	Insight into hepatocellular carcinogenesis at transcriptome level				
	by comparing gene expression profiles of hepatocellular carcinoma				
	with those of corresponding noncancerous liver				
	Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)				
	21625106				
	Contact: Zeguang Han				
	Chinese National Human Genome Center at Shanghai				
	351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai				
	201203, P. R. China				
	Tel: 86-21-50801919(ex.45)				
	Fax: 86-21-50801922				
	Email: hanzg@chgc.sh.cn				
	This clone is available at CHGC in Shanghai.				
FEATURES	Location/Qualifiers				
SOURCE	1..679				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone="GLCBD03"				
	/clone_id="GLC"				
	/tissue_type="corresponding non cancerous liver tissue"				
	/dev_stage="Adult"				
	/lab_host="SOLR"				
	/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:				

BASE COUNT	199 a	151 c	206 g	116 t	7 others
ORIGIN					
Alignment Scores:					
Pred. No.:		3.55e-66			679
Score:		889.00			174
Percent Similarity:		95.19%			Conservative: 4
Best Local Similarity:		93.05%			Mismatches: 19
Query Match:		60.03%			Indels: 10
DB:		10			Gaps: 0
US-09-513-365A-1 (1-280) x AV647917 (1-679)					
QY	1	Metleunglygninginglinginglnleu	yserserlaalaaleu	thrglyluar	20
DB	117	ATGTTAGGGCAGCAGCAGCAGCACTGACTGCTGCGGCTGCGCTCGACCGGGAGCGG			176
QY	21	SerargleuethrCysTyrValGlnAspTyrleuGluCysValgluserleuProHis			40
DB	177	AGCGGGCGCTCACCCTGCTAGCTCAGGACACCTTGCTGCGGAGTCCCTCCAC			236
QY	41	AspmetGlnArgAsnValSerValleuArgGluLeuAspAsnLysTyrGlnGluThrleu			60
DB	237	GACATGCAGAGAGAACGCTGTCTGTCTGCGGAGACCTGGACACCAATATCAAGAACTTA			296
QY	61	LysgluIleaspaspValTyrGluLysTyrLysLysgluAspAspLeuAsnGlnLysLys			80
DB	297	AAGGAATATGATGATGCTCTACGAAAAATATTAAGAAAGATGATTTAAACAGAGAA			356
QY	81	ArgleunginglnleuGlnArgAlaIleuIleAsnserginglnleuGlyAspGluLys			100
DB	357	CGTCTACAGCAGCTCTCTCCAGAGACACTAATTAATAGTCAAGATGGGAGATGA			416
QY	101	IleGlnIleValThrGlnMetleungluLeuValGlnAsnArgAlaArgGlnMetGluLeu			120
DB	417	ATCAGATGGTTACACCAATCTCTCGAATGGTNNGAAATCGCCACCAATGANGTA			476
QY	121	HisserGlnCysPheGlnAspProAlaGluSerGluArgAlaSerAspLysAlaLysMet			140
DB	477	CACCTACAGTGTTTNCAGATCCGCTGTAAGTGAAGCAGCCTCAGATTAAGCAGNATG			536
QY	141	AspserSerGlnProGluArgSerSerArgArgProArgArgGlnArgThrSerGluSer			160
DB	537	GATTTCACCCACCAAGAAAGATCTTCAAGAAAGACCCCGAGCAGCGGACCAAGCAAGC			596
QY	161	ArgAspIleuGlnHisMetAlaAsnGlyIleGluLysPysAspAspGlnProProGlu			180
DB	597	CGTGATTTATGTGCATGTGCAAAATGGGATTAAGACCTGTGATGATCAGGCACTTAAGAA			656
QY	181	LysLysSerLysSerAlaLys			187
DB	657	AAGAAATCCAAAGTCAAGCAAG			677
RESULT 14					
LOCUS	BE839427/c		564 bp	mrna	linear EST 22-SEP-2000
DEFINITION	RC3-FN0143-190700-022-d12 FN0143				Homo sapiens cDNA, mRNA sequence.
ACCESSION	BE839427				
VERSION	BE839427.1				GI:10271805
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE					
AUTHORS	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 564)				
	Dias Neto, E., da Silva, W. Jr., Zago, M. A., Boridin, S., Costa, F. F., Nagai, M. A., Carvalho, A. F., Matsukuma, A., Bata, G. S., Simpson, D. H., Goldman, G. H., de Oliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentiani, R. R., Reis, L. F., de Souza, S. J. and Simpson, A. J.				
	Shotgun sequencing of the human transcriptome with ORF expressed				







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 20, 2003, 01:27:39 ; Search time 241 Seconds

(without alignments)  
2616.430 Million cell updates/sec

Title: US-09-513-365A-1  
1481  
Sequence: 1 MCGGQQQQLVSSAALLTGER.....DNEKTMKSTKTKKDRSR 280

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Ygapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+P2n.model -DEV=xlP  
-O=/cgn2\_1/USPTO.spool/US09513365/funat\_16062003\_124742\_16434/app-query.fasta.1.455  
-DB=N\_Geneseq\_101002 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=Dlosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MIMER=0 -MATEL=2000000000  
-USER=US09513365\_ECGN1.1\_1366\_etunat\_16062003\_124742\_16434 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEOBJECT -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV.TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq\_101002: \*  
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: \*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: \*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT: \*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: \*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: \*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: \*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: \*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: \*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: \*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: \*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: \*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: \*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: \*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: \*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: \*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: \*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: \*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: \*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: \*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: \*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: \*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: \*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: \*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1481	100.0	840	20	Human INGL1 encodi
2	1481	100.0	1078	20	Human INGL1 encodi
3	1481	100.0	1080	22	Tumour suppressor
4	1345	90.8	1153	21	Human tumour suppressor
5	874.5	59.0	2817	21	Murine p37ING1 cod
6	853.5	57.6	911	21	Human p37ING1 codi
7	853.5	57.6	1533	22	Nucleotide sequenc
8	853.5	57.6	2897	24	Human inhibitor of
9	745	50.3	1835	21	Murine ingl common
10	728	49.2	1143	22	Nucleotide sequenc
11	728	49.2	1902	18	Tumour suppressor
12	728	49.2	1902	19	Partial INGL part
13	728	49.2	2061	18	Tumour suppressor
14	728	49.2	2061	19	Human INGL full-le
15	728	49.2	2061	20	Nucleotide sequenc
16	728	49.2	2886	24	Human inhibitor of
17	728	49.2	8487	22	Human musculoskele
18	604	40.8	742	22	Nucleotide sequenc
19	602	40.6	857	22	Nucleotide sequenc
20	572	38.6	346	21	Human secreted pro
21	480	32.4	693	24	Oligonucleotide fo
22	480	32.4	693	24	Oligonucleotide fo
23	466	31.5	693	24	Oligonucleotide fo
24	466	31.5	693	24	Oligonucleotide fo
25	455	30.7	1705	22	Human polynucleoti
26	447	30.2	1772	23	Human polynucleoti
27	447	30.2	1781	21	Drosophila melanog
28	446.5	30.1	1465	21	p33 tumour suppress
29	446.5	30.1	1611	22	cDNA encoding a hu
30	446.5	30.1	1611	22	Human tumour suppressor
31	436.5	29.5	987	23	Human polynucleoti
32	437.5	28.9	1365	24	Drosophila melanog
33	427.5	28.9	1108	22	Human tumour suppressor
34	424.5	28.7	958	22	Cell cycle protein
35	406	27.4	1465	22	Human polynucleoti
36	400	27.0	699	21	Human polynucleoti
37	396.5	26.8	678	21	Human polynucleoti
38	392.5	26.5	807	22	Human polynucleoti
39	392.5	26.5	807	22	Human polynucleoti
40	387.5	26.2	3914	23	Cell cycle protein
41	384.5	26.0	958	22	Drosophila melanog
42	362	24.4	683	24	Cell cycle protein
43	362	24.4	683	24	Oligonucleotide fo
44	359	24.2	1807	22	Oligonucleotide fo
45	359	24.2	1864	22	cDNA encoding huma

## ALIGNMENTS

RESULT 1  
AAZ08595  
ID: AAZ08595 standard; cDNA: 840 BP.

AAZ08595; 18-OCT-1999 (first entry)

DE Human INGL1 encoding cDNA.

KW Human: TSC403; INGL1; diagnosis; lung cancer; cell cycle; regulation;

KW cell proliferation; cell aging; apoptosis; tumour suppressor; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

XX FT 1..840

XX FT CDS /\*tag= a





\_\_\_\_\_

QY

1 Met



```

OY 11leasnsrerglnlueuglyaspglulysllleglnlvalthrnglnmetleugluene 110
DB 427 ATTATATGTCAGATATGGAGATGATAAATAATACAGATGTGTACACAATCTCGAATG 486
OY 111 Valgluasnargalaarglnmetglulueuhsrerglnncysphaglnaspralaglu 130
DB 487 GTGGAAATCGGGCAGACCAATGAGTTACACACAGTGTTCACCAAGATCCGCTGAA 546
OY 131 Sercluaralaseraplysalaalysmetaspserserglnprogluarargserarg 150
DB 547 AGTCAGAGAGCCCTCAGATTAAGCAAAATGATTCAGCCCAACAGAAAGATCTTCAAGA 606
OY 151 ArproarararglnarargthrsrerglnserargaspleucysHsmetalasnglyle 170
DB 607 AGACCCCGCAGCGAGCGAGCAGTGAAGCCGTAATTTATGTCAATGCAATGGGCAAT 666
OY 171 Gluaspcysaspaprglnproprolysglulyslyserserlalaalysalyslys 190
DB 667 GAAGACTGTGATGTATCAGCCACCTAAAGAAAGAAATCCAGACGCAAGAAAGAA 726
OY 191 Argserlysalaalysglnluarglualaserprovalgluphealialaspproasn 210
DB 727 CGCTCCAGGCGCAGGAGGAAAGGAAAGCTTCACTGTGAGTTGCAATGATCTTAT 786
OY 211 Gluprothrrtyrcysleucysasnlnvalserlyrglylmetlleglucysaspasn 230
DB 787 GAACCTACATCTCTCTATGCAACCAAGCTCTTATGGGAGATGATGATGATGATGAT 846
OY 231 GluIngcysprolelguttrpPheHsPheSerCysValSerLeuthrrtyrlyspProlys 250
DB 847 GAACAGTGTCCAAATGGAATGTTTCACTTTCATGTGTTTCACTTACCTATAACCAAG 906
OY 251 Glyysrrprrtyrcysprolyscysarglyaspasnclulysstnmeaspysserthr 270
DB 907 GGGAAATGATTTCCCAAAAGTCAGGGGAGATATAGAAACAAATGACAAAAGTACT 966
OY 271 GlulysThrlyslasparargserarg 280
DB 967 GAAGAAGCAAAAAGATAGAGATCGAGC 996

RESULT 5
AAA53790
ID AAA53790 standard; cDNA: 2817 BP.
XX
AC AAA53790;
XX
DT 19-DEC-2000 (first entry)
XX
DE Murine P37ING1 coding sequence.
XX
KW p53; tumour; cancer; detection; antibody; hybridisation; exon 1b;
KW INGI; lngl; p37ING1; p37ING1; oncogene; gene therapy; diagnosis;
KW proliferation disorder; transformation; transformed cell; mouse; ds.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 847..1686
FT /tag= a
FT /product= P37ING1 polypeptide
PN WO200046370-A1.
XX
PD 10-AUG-2000.
XX
PF 04-FEB-2000; 2000WO-US02959.
XX
PR 04-FEB-1999; 99US-0118941.
XX
PA (UNIT ) UNIV ILLINOIS FOUND.
PI Gudkov A, Zeremski M, Gurova KV, Grigorian IA;

```

```

XX WPI: 2000-491278/43.
DR P-PSDB: AAY97242.
XX
PR Detecting nucleic acid encoding exon 1b of lngl, useful for diagnosing
PR and treating cancer, comprises contacting sample with isolated nucleic
PR acid comprising sequence of exon 1b and detecting hybridized products
XX
PS Disclosure: Fig 11; 134pp; English.
XX
CC Mutations in or loss of the p53 gene occur in more than 50% of
CC human tumors and tumour cell lines, but functional inactivation of
CC the p53 pathway occurs in a much larger proportion of tumours. In
CC many cases the mechanism of functional inactivation of the p53 gene
CC remains unknown but p53 has been found to act in cooperation with
CC INGI. Functional cooperation between INGI and p53 suggested that
CC INGI encoded a tumour suppressor protein that functioned within the
CC p53 pathway. This data suggested a possible role for INGI in head
CC and neck cancers and chromosomal location of the INGI placed it
CC within a region that is frequently rearranged in head and neck
CC cancers. Large scale analysis of tumours involving INGI has not
CC revealed mutations in INGI nor significant variations in its
CC expression suggesting that INGI was not a useful gene to study in
CC cancer etiology. However, alternative initiation exons of the lngl
CC gene, each having their own promoter have been discovered.
CC Expression of one promoter (1a) produces a protein identical to
CC INGI. Expression of a second promoter (1b) produces a protein having
CC an identical C-terminal fragment to INGI but an additional 104
CC N-terminal amino acids. The newly discovered protein has been
CC designated p37ING1 (wild type: p37ING1). p37ING1 has the
CC characteristics of an oncogene. When overexpressed in cells (even
CC those expressing wild type p53) p37ING1 is able to cause
CC proliferation or transformation of those cells. Thus detecting a
CC nucleic acid encoding exon 1b of lngl by hybridisation with an
CC isolated nucleic acid having the sequence of exon 1b of lngl
CC or its antisense sequence can identify individuals expressing the
CC oncogenic form of lngl. Novel peptide sequences taken from the 104
CC N-terminal peptide of p37ING1 can also be used to raise antibodies
CC that can also be used in detection methods for the p37ING1 variant.
CC The polypeptides may be useful in gene therapy for treatment of cell
CC proliferation disorders, especially cancers and for diagnosing and
CC studying cancers.
XX
SQ Sequence 2817 BP; 670 A; 711 C; 846 G; 590 T; 0 other:
XX
Alignment Scores:
Pred. No.: 7,04e-74 Length: 2817
Score: 874.50 Matches: 165
Percent Similarity: 77.41% Conservative: 44
Best Local Similarity: 61.11% Mismatches: 50
Query Match: 59.05% Indels: 11
DB: 21 Gaps: 3
XX
US-09-513-365A-1 (1-280) x AAA53790 (1-2817)
OY 18 GlyluargserargleuLeuThrcysrtyrValgluaspyrleugluValgluser 37
DB 865 GGGGAGCAGATCCACCTGTGTGAC--TATGGGAGGATTCCTGACTCAATGACGCA 921
OY 38 LeuproHIsaspmetgluaraspsnValserValleuAAGglulueaspaslnlystgln 57
DB 922 CTGCCTTTCAGACTGCAGAGACGCTCGCTGATGGCGGAGATGAGACGCCAATACCAA 981
OY 58 GluthrleuLysglulileaspaspValtyrclulyslyrlysgluaspaspleuasn 77
DB 982 GAGATCTCGAAGAGAGCTGAGACGACTATGAGAAGTCAACAGGAGACAGACGGCACC 1041
OY 78 GlulyslysargleuglnleuLeuenglnrAlaLeuIleasnsrerglnluengly 97
DB 1042 CAGAGCGCGGGTACTGCTCACTCCAGAGGGCGCTGATCCGACGACGAGCTAGGC 1101
OY 98 AspgluylleaglilleValthrnglnmetleugluLeuValgluasnaralargln 117

```

```

Db 1102 GATGAGAGATTCAGATCGTACGTGAGCTGGTGGAGAACCCGACGACAG 1161
Qy 118 MetGluDeuHisserGlnCysPhe-----GlnAspProAlaGlnSerGluArgAla 134
Db 1162 GTGGACAGTACGTCGAGCTCTTGAAGACACACGACATCACTGACGCTGGTGC 1221
Qy 135 SerAspLysAlaLysMetAspSerSerGlnProGluArg----- 147
Db 1222 AGCGGACAGCGGGCGGACGACAAAGTCAGACGCTCCTCCTGCGGACCTTCCATC 1281
Qy 148 SerSerArgArgProArgArgGlnArgThrSerGlnSerArgAspLeuGlnSerMetAla 167
Db 1282 AATTAACAGAGGCTCCAGAGGACGACGAAACAAATGAGTGAAGAAAGCCGTCGAAATAT 1341
Qy 168 AsnGly11LeuLysAspCysAspAspGlnProProLysGluLysSerLysSerAlaLys 187
Db 1342 CACGACCATGATGATCACCACCTCAGAACGCCCAAGAGAAAGAAAGAAACCTCAAG 1401
Qy 188 LysLysLysArgSerLysAlaLysGlnGlnArgGlnAlaSerProValGluPheAlaIle 207
Db 1402 AAGAGAAACGCTCCAGGCGCAAGAGGAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGG 1461
Qy 208 AspProAsnGlnProThrTrpCysLeuCysAsnGlnValSerTrpGlyMetIleGly 227
Db 1462 GACCCACAGAGCCGACGCTGCTGTGCAACAGGCTCTCTACGGGGAGATGATGAGG 1521
Qy 228 CysAspAsnGlnLysCysProLysGlnArgPheHisPheSerCysValSerLeuThrTyr 247
Db 1522 TGTGACAAAGCAAGATGCCCCCATGCGAGTGTCCACTTCTCCGCGGGGCTCAACCAT 1581
Qy 248 LysProLysGlnLysTrpTyrCysProLysCysArgGlnLysAsnGlnLysMetAsp 267
Db 1582 AAACCAAAAGGCAAGTGTACTGCCCCCAAGTGCCTGGGAGAGCGAAGAACCATGAC 1641
Qy 268 LysSerThrGlnLysThrLysLysAspArg 277
Db 1642 AAAGCCCTGGAGAGATCCAGAAAGAGAGG 1671

RESULT 6
AAAS3792
ID AAAS3792 standard; cDNA; 911 BP.
XX
AC AAAS3792;
XX
DT 19-DEC-2000 (first entry)
XX
DE Human P37ING1 coding sequence.
XX
KW p53; tumour; cancer; detection; antibody; hybridisation; exon 1b;
KW ING1; Ing1; p33ING1; p37ING1; oncogene; gene therapy; diagnosis;
KW Proliferation disorder; transformation; transformed cell; human; ds.
XX
OS Homo sapiens.
XX
PN WO200046370-A1.
XX
PD 10-AUG-2000.
XX
PF 04-FEB-2000; 2000WO-US02959.
XX
PR 04-FEB-1999; 99US-0118941.
XX
PA (UNIT ) UNIV ILLINOIS FOUND.
XX
PI Gudkov A, Zeremski M, Gurova KV, Griorian IA;
XX
DR P-FSDB; AAY97244.
XX
WPI: 2000-491278/43.
XX
PT Detecting nucleic acid encoding exon 1b of Ing1, useful for diagnosing
PT acid comprising sequence of exon 1b and detecting hybridized products
XX

```

```

PS Disclosure; Page 122-123; 134pp; English.
XX
CC Mutations in or loss of the p53 gene occur in more than 50% of
CC human tumours and tumor cell lines, but functional inactivation of
CC the p53 pathway occurs in a much larger proportion of tumours. In
CC many cases the mechanism of functional inactivation of the p53 gene
CC remains unknown but p53 has been found to act in cooperation with
CC ING1. Functional cooperation between ING1 and p53 suggested that
CC ING1 encoded a tumor suppressor protein that functioned within the
CC p53 pathway. This data suggested a possible role for ING1 in head
CC and neck cancers and chromosomal location of the ING1 placed it
CC within a region that is frequently rearranged in head and neck
CC cancers. Large scale analysis of tumours involving ING1 has not
CC revealed mutations in ING1 nor significant variations in its
CC expression suggesting that ING1 was not a useful gene to study in
CC cancer etiology. However, alternative initiation exons of the Ing1
CC gene, each having their own promoter have been discovered.
CC Expression of one promoter (1a) produces a protein identical to
CC ING1. Expression of a second promoter (1b) produces a protein having
CC an identical C-terminal fragment to ING1 but an additional 104
CC N-terminal amino acids. The newly discovered protein has been
CC designated p37ING1 (wild type: p33ING1). p37ING1 has the
CC characteristics of an oncogene. When overexpressed in cells (even
CC those expressing wild type p53) p37ING1 is able to cause
CC proliferation or transformation of those cells. Thus detecting a
CC nucleic acid encoding exon 1b of Ing1 by hybridisation with an
CC isolated nucleic acid having the sequence of exon 1b of Ing1
CC or its antisense sequence can identify individuals expressing the
CC oncogenic form of Ing1. Novel peptide sequences taken from the 104
CC N-terminal peptide of p37ING1 can also be used to raise antibodies
CC that can also be used in detection methods for the p37ING1 variant.
CC The polypeptides may be useful in gene therapy for treatment of cell
CC proliferation disorders, especially cancers and for diagnosing and
CC studying cancers.
XX
SQ Sequence 911 BP; 228 A; 255 C; 305 G; 123 T; 0 other.

Alignment Scores:
Pred. No.: 1.65e-72 Length: 911
Score: 853.50 Matches: 161
Percent Similarity: 75.74% Conservative: 45
Best Local Similarity: 59.19% Mismatches: 51
Query Match: 57.63% Indels: 15
DB: Gaps: 3

US-09-513-365A-1 (1-280) x AAAS3792 (1-911)
Qy 18 GlyGluArgSerArgLeuLeuThrCysTyrValGlnAspTyrLeuGlnCysValGlnSer 37
Db 63 GGGGAGCAGCTCCACCTGGTGAAC--TATGTGAGAGACTACCTGGACTCCTCATCGATCC 119
Qy 38 LeuProHisAspMetGlnArgAsnValSerValLeuArgGlnLeuAsnAspLeuTyrGln 57
Db 120 CTGCCTTTCACCTTGAGAGAAATGCTGCTGATGCGGAGATGACGCGGAATATACCA 179
Qy 58 GlnThrLeuLysGlnLeuAspAspValTyrGlnLysTyrLysGlnAspLeuAsn 77
Db 180 GAGATCCTGAAGAGAGCTAGAGAGAGTGTACAGCGCTTCACTGCGGAGACAGCGGGCG 239
Qy 78 GlnLysLysArgLeuGlnGlnLeuLeuGlnArgAlaLeuLeuAsnSerGlnGlnLeuGly 97
Db 240 CAGAGGCGGCGAGTGTGACACTGTGTGACGCGGCGCTGTATCCGACCCAGGAGCTGGGC 299
Qy 98 AspGlnLysIleGlnIleValThrGlnMetLeuGlnLeuValGlnAsnArgAlaArgGln 117
Db 300 GACGAGAGATTCAGATCGTGAAGAGATGTGAGAGTGTGAGAAACCGACGCGGACG 359
Qy 118 MetGluDeuHisserGlnCysPheGlnAspProAlaGlu----- 130
Db 360 GTGGACAGCCAGCTGAGTGTGAGAGGCGACAGAGCTGGGCGGACACAGTGGGAC 419
Qy 131 -----SerGluArgAlaSerAspLysAlaLysMetAspSerSerGlnPro 145

```

DB 420 AGCGCAGAGTTGGCGCGACAGGCCCAATGCGCATGCGGTAGCGCACTGTGACAAAGCCC 479  
QY 146 GluArgSerSerArgProArgArgGlnArgThrSerGlnuserArgAspLeucylshs 165  
DB 480 -----AACGACGACCGCTCACGGGCGACGCAACAGAACCGTGAGAACCGCTCC 533  
QY 166 MetAlaAsnGlyIleGluAspCysAspAspGlnProPheGlyIleGlySerLysSer 185  
DB 534 AGCAACACGACGACGACGAGCGGCGCTCGGCGACACCCCAAGAGAAAGAGCCAGAAC 593  
QY 186 AlalysIlyslYslYslYslYslYslYslYslYslYslYslYslYslYslYslYslYsl 205  
DB 594 TCCAG 653  
QY 206 AlaIleAspProAsnGluProThrThrThrCysLeucylYslYslYslYslYslYslYsl 225  
DB 654 CCCATCGACCCCAAG 713  
QY 226 IleGlyCysAspAsnGluGlnCysProIleGluThrPheHisPheSerCysValSerLeu 245  
DB 714 ATCGGCTCGACGACGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 773  
QY 246 ThrTrpTrpProLysGlyIleYslYslYslYslYslYslYslYslYslYslYslYsl 265  
DB 774 AATCATAAACCAAG 833  
QY 266 MetAspLysSerThrGluIleYslYslYslYslYslYslYslYslYslYslYslYsl 277  
DB 834 ATGACAAAGCCCTGAGAGAAATCCAAAAAGAGAG 869

RESULT 7  
AAH28478  
ID AAH28478 standard; DNA; 1533 BP.  
AC AAH28478;  
AC AAH28478;  
DT 17-SEP-2001 (first entry)  
XX Nucleotide sequence of a human cancer associated antigen.  
DE Nucleotide sequence of a human cancer associated antigen.  
XX  
XX Cancer associated antigen; INGI; tumour suppressor; cancer; vaccine; ss.  
XX  
XX Homo sapiens.  
OS  
XX  
XX Key Location/Qualifiers  
FH CDS 451..1290  
FT  
FT /\*tag= a  
FT /transl\_except= "(pos: 817..819, aa: Val)"  
FT /transl\_except= "(pos: 835..837, aa: Val)"  
FT /transl\_except= "(pos: 853..855, aa: Asn)"  
FT /transl\_except= "(pos: 859..861, aa: Asp)"  
FT /transl\_except= "(pos: 865..867, aa: Val)"  
FT /transl\_except= "(pos: 874..876, aa: Ser)"  
FT /product= "cancer associated antigen"

MO200147959-A2.  
PN  
XX  
XX 05-JUL-2001.  
PD  
XX  
XX 29-NOV-2000; 2000WO-US42334.  
PF  
XX  
XX 30-NOV-1999; 99US-0451739.  
PR 24-OCT-2000; 2000US-0602362.  
XX  
XX  
XX (LUDWIG INST CANCER RES.  
PA (SLOK ) SLOAN KETTERING INST CANCER RES.  
PA (CORR ) CORNELL RES FOUND INC.  
PI Jager D, Stockert E, Scanlan M, Knuth A, Old L, Gure A, Chen Y;  
XX  
XX WPI: 2001-441706/47.  
XX P-PSDB: AAB84698.

PT Isolated cancer associated nucleic acid molecule identified by SEREX  
PT (serological identification of antigens by recombinant expression  
PT cloning) technique, useful in nucleic acid based therapies to treat  
PT cancer -  
XX  
XX PS Claim 1; Page 43-44; 62pp; English.  
XX  
XX The present sequence encodes a human cancer associated antigen.  
CC The sequence is a variant of the INGI gene, which is a tumour  
CC suppressor gene candidate. The cancer associated antigen polynucleotides  
CC and polypeptides are useful for screening for the possible presence of  
CC a pathological condition in a subject such as cancer. The cancer  
CC associated antigen polypeptides are useful for producing vaccines.

XX  
SQ Sequence 1533 BP; 336 A; 431 C; 521 G; 244 T; 1 other;  
Alignment Scores:  
Pred. No.: 3,27e-72 Length: 1533  
Score: 853.50 Matches: 160  
Percent Similarity: 76.30% Conservative: 46  
Best Local Similarity: 59.26% Mismatches: 53  
Query Match: 57.63% Indels: 11  
DB: 22 Gaps: 3

US-09-513-365A-1 (1-280) x AAH28478 (1-1533)  
QY 18 GlyIuArgSerArgLeuLeuThrCysTrpYslYslYslYslYslYslYslYslYslYslYsl 37  
DB 469 GGGGAGCAGCTCCACCTGCTGTAAC--TATGTGAGAGAGACTACTGCTGCTCCAGATCC 525  
QY 38 LeuProHisAspMetGlnArgAsnValSerValIleArgGluLeuAspAsnLysTrpGln 57  
DB 526 CTGCTTTCGACTTCAGAGAAATCTTCGCTGATCGCGGAGATCGACGCGAAATACCA 585  
QY 58 GlnThrLeuLysGluIleAspAspValYslYslYslYslYslYslYslYslYslYslYsl 77  
DB 586 GAGATCTGAAAG 645  
QY 78 GlnLysLysArgGluGlnLeuLeuGlnArgGlnArgGlnArgGlnArgGlnArgGlnArg 97  
DB 646 CAGAGCGCGGATCTCCACTGTGTGAGCGCGCGGCTGATCCGACGAGAGAGCTGGGC 705  
QY 98 AspGluLysIleGlnIleValThrGlnMetLeuGlnLeuValGlnAsnArgGlnArgGln 117  
DB 706 GACGAGAAATCAATGCTGTGAGCCAGATGCTGTGAGAGAGAGAGAGAGAGAGAGAGAG 765  
QY 118 MetGluLeuHisSerGlnCysPheGlnAspProAlaGlu--SerGluArgGlnAsp 136  
DB 766 GTGACAGACGAGCTGAGAGCTGTTCGAGCGCAGAGAGAGCTGGGAGACACAGCGGCAAC 825  
QY 137 LysAlaLysMetAspSerSerGlnPro-----GluArg 147  
DB 826 AGCGGCAAGGCTGGCGGACGACGCCCAAGGCGGAGCGGCGGAGCTGACCAAGCC 885  
QY 148 SerSerArgArgProArgArgGlnArgThrSerGlnuserArgAspLeucylshsMeta 167  
DB 886 AACACAGAGCGCTCACGGGCGACGCAACAGAGAACCGTGAGAACCGCTCCAGCAAC 945  
QY 168 AsnGlyIleGluAspCysAspAspGlnProPheGlyIleGlySerLysSerLys 187  
DB 946 CACGACACGACGAGCGGCGCTCGGCGACACCCCAAGAGAGAGAGAGAGAGAGAGAG 1005  
QY 188 LysLysLysArgSerLysAlaLysGlnGlnArgGlnArgGlnArgGlnArgGlnArgGln 207  
DB 1006 AAGAGAGAGCGCTCCAGAGCGGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1065  
QY 208 AspProAsnGluProThrThrThrCysLeucylYslYslYslYslYslYslYslYslYsl 227  
DB 1066 GACCCCAAG 1125  
QY 228 CysAspAsnGluGlnCysProIleGluThrPheHisPheSerCysValSerLeuThrTrp 247  
DB 1126 TCGACAAAG 1185





```

FT      /transl_except= "(pos: 124..126, aa: Pro, Ala)"
FT      /product= "Cancer associated antigen"
XX      MO200147959-A2.
XX      05-JUL-2001.
XX      29-NOV-2000; 2000WO-US42334.
XX      30-NOV-1999; 99US-0451739.
XX      24-OCT-2000; 2000US-0602362.
XX      (LUDW-) LUDWIG INST CANCER RES.
XX      (SLOK ) SLOAN KETTERING INST CANCER RES.
XX      (CORR ) CORNELL RES FOUND INC.
XX      Jager D, Stockert E, Scanlan M, Knuth A, Old L, Gure A, Chen Y;
XX      WPI; 2001-441706/47.
XX      P-PSDB; AAB84697.
XX      Isolated cancer associated nucleic acid molecule identified by SEREX
XX      (serological identification of antigens by recombinant expression
XX      cloning) technique, useful in nucleic acid based therapies to treat
XX      cancer.
XX      Example 4; Page 44; 62pp; English.
XX      The present sequence encodes a human cancer associated antigen.
XX      The sequence is the wildtype of the INGI gene, which is a tumour
XX      suppressor gene candidate. The cancer associated antigen polypeptides
XX      and polypeptides are useful for screening for the possible presence of
XX      a pathological condition in a subject such as cancer. The cancer
XX      associated antigen polypeptides are useful for producing vaccines.
XX      Sequence 1143 BP; 289 A; 291 C; 366 G; 197 T; 0 other;
SQ
Alignment Scores:
Pred. No.:      2,37e-60      Length:      1143
Score:          728.00      Matches:      136
Percent Similarity: 74.79%      Conservative: 39
Best Local Similarity: 58.12%      Mismatches: 45
Query Match:    49.16%      Indels:      14
DB:             22      Gaps:        2
US-09-513-365a-1 (1-280) x AAB28479 (1-1143)
QY      56 TyGInGluThrLeuysGluIleaspAspValTyrGluTyrIlySylsGluaspAsp 75
DB      190 TGGAAACAGATCTGTAAGAGAGCTAGACAGAGTGTCTACAGAGCGCTTCACTCGCGAGACAGAC 249
QY      76 LeuAsnGlnIlySylsArgLeuGlnGlnLeuLeuGlnArgAlaLeuIleAsnSerGlnGlu 95
DB      250 GGGGCGCAGAAAGCGCGAGTGTGCACTGTGTGTGAGCGCGCGCTGATGCCAGCAGCAGAG 309
QY      96 LeuGlyAspGluIlySylsIleGlnIleValThrGlnMetLeuGlnLeuValGluAsnArgAla 115
DB      310 CTGGGGGAGCAGAAATCCAGATCGTGTGAGCCACAGATGTGTGAGTGAGGAGAACCGCAGC 369
QY      116 ArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGlu----- 130
DB      370 CGGCGAGTGGACAGCAGCTGTGAGCTGTGTGAGCGCAGCAGCAGCGACACAGTG 429
QY      131 -----SerGluArgAlaSerAspIlySylsMetAspSerSer 143
DB      430 GGCACACAGCGCAAGTGTGGCGCGAGAGCCCAATGCGCATGCGGTAGCGCAGTCTGAC 489
QY      144 GlnProGluArgSerSerArgArgProArgArgGlnArgThrSerGlnSerArgAspLeu 163
DB      490 AAGCCCG-----AAGCAGCAGCGCTCAGCGCGCAGGCAACAGCAGAACCGTGAGAAC 543
QY      164 CysHisMetAlaAsnGlyIleGluAspCysAspAspGlnProProIlySylsGluIlySylSer 183

```

```

DB      544 GCGTCCAGCAACAGCAGCAGCAGCGCGCTCGGGCAGACCCCAAGAGAAAGAGCC 603
QY      184 LysSerAlaIlySylsIlySylsArgSerIlyAlaIlyGlnGlnArgGluAlaSerProVal 203
DB      604 AAGACCTCCAGAAAGAAAGAGCGCTCCAAAGGCCAAGCGGAGCGAGGGCTCCCTGCC 663
QY      204 GluPheAlaIleaspProAsnGluProThrTyrCysLeuGlnCysAsnGlnValSerTyrGly 223
DB      664 GACCTCCCATGCGACCCCAAGAACCCAGCTACTGTGTGTGCAACAGAGTCTCTATGGG 723
QY      224 GluMetIleGlyCysAspAsnGlnGlnCysProIleuIleuThrPheHisPheSerCysVal 243
DB      724 GAGATGATGCGCTCGCGCAACAGCAGAGTGTCCCATGAGTTCACCTCTCGTGGTG 783
QY      244 SerLeuThrTyrIlySylsProIlySylsTyrTyrCysProIlyCysArgGlyAspAsnGlu 263
DB      784 GGGCTCAATCATTAACCAAGGCGAGTGTACTGTCTCCAAAGTCCGGGGGAGAACGAG 843
QY      264 LysThrMetAspIlySylsSerThrGlnIlySylsIlySylsAspArg 277
DB      844 AAGACATGAGCAAAAGCCCTGAGAAATCCAAAAAAGAGAG 885
RESULT 11
AAT69651
ID      AAT69651 standard; cDNA; 1902 BP.
AC      AAT69651;
XX      27-AUG-1997 (first entry)
XX      Tumour suppressor gene INGI partial cDNA.
XX      Tumour suppressor gene; INGI: p33ING1; breast cancer; brain cancer;
XX      diagnosis; gene therapy; ss.
XX      Homo sapiens.
XX      Key      Location/Qualifiers
XX      CDS      109..741
XX      FT      /*tag= a
XX      PN      W09721809-A1.
XX      PD      19-JUN-1997.
XX      PE      06-DEC-1996; 96WO-CA00819.
XX      PR      15-NOV-1996; 96US-0751230.
XX      PR      08-DEC-1995; 95US-0569721.
XX      PA      (UYTE-) UNIV TECHNOLOGIES INT INC.
XX      PI      Garkavstev I, Ribowol K;
XX      DR      WPI; 1997-332781/30.
XX      DR      P-PSDB; AAW19118.
XX      PT      Isolated tumour suppressor gene, INGI - useful to develop products
XX      for inhibiting or increasing cell proliferation, in particular for
XX      treatment or diagnosis of cancer
XX      Claim 1; Fig 2; 63pp; English.
XX      A partial cDNA clone (AAT69651), designated INGI, codes for a novel
XX      tumour suppressor protein p33ING1 (AAW19118) that is a potent
XX      inhibitor of cell growth. It was isolated by subtractive
XX      hybridisation between normal mammary and transformed epithelial
XX      cDNAs. Isolation of an antisense INGI cDNA insert that caused
XX      increased cell proliferation, and use of the insert to screen
XX      normal human fibroblast and HeLa cDNA libraries. A complete INGI
XX      sequence (AAT69652) was obt'd. by 5'RACE. INGI is localised to the
XX      19q33-34 chromosome region, to which a number of human cancers have
XX      been mapped. INGI nucleic acids can be used in the diagnosis of

```



CC breast cancer; a decreased level of INGI mRNA indicates cancerous  
CC cells. They can also be used in gene therapy methods to block the  
CC proliferation of cancer cells.

XX Sequence 1902 BP; 574 A; 391 C; 461 G; 476 T; 0 other;

# Alignment Scores:

Pred. No.:	Length:	1902
Score:	Matches:	136
Percent Similarity:	Conservative:	39
Best Local Similarity:	Mismatches:	45
Query Match:	Indels:	14
	Gaps:	2

US-09-513-365a-1 (1-280) x AAT69651 (1-1902)

```

OY 56 TyrcInGIuThrLeuLysGluIleAspAspValTyrcGluTyrcLysLysLysAsp 75
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 31 TGGAAACAGATCCGAGAGAGCTAGACGAGTGTACAGAGCGCTCAGTCCGAGACAGAC 90
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 76 LeuAsnGIuLysArgLeuGIuInLeuLeuGIuInArgAlaLeuIleAsnSerGIu 95
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 91 GGGCGCAGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 150
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 96 LeuGIAspGIuLysIleGIuInLeuValThrGIuInLeuGIuValGIuAsnArgAla 115
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 151 CTGGCGCAGCAGAGATCCAGATGTCAGCCAGATGTGTGAGCTGTGTGAGACACCGACG 210
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 116 ArgGIuMetGIuLeuHisSerGIuCysPheGIuAspProAlaGIu----- 130
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 211 CGGAGGTGACAGCCAGCTGTGAGAGCTGTGAGAGCGCAGCAGAGCTGGCGACACAGTG 270
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 131 -----SerGIuArgAlaSerAspLysAlaLysMetAspSer 143
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 271 GGCAACAGCGCAGAGTGGCGGAGAGCGCCAGCGCATGGGAGCGCAGCTGTGAC 330
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 144 GIuProGIuArgSerArgArgProArgArgIuArgIuArgIuSerGIuSerArgAspLeu 163
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 331 AAGGCC-----AACAGCAGAGCGCTCAGCGGCGAGCGCAACAGAGAACCGTGAAC 384
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 164 CysHisMetAlaAsnGIuIleGIuAspCysAspAspGIuProPolysGIuLysSer 183
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 385 GCGTCACGACACGACAGAGAGCGGCGCTCGGCGACACCCAGAGAGAGAGAGGCC 444
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 184 LysSerAlaLysLysLysArgSerLysAlaLysGIuArgGIuAlaSerProVal 203
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 445 AAGACCTCCAGAGAGAGAGCGCTCCAGAGCGCAGCGAGCGCTCCCGCTGCC 504
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 204 GIuPheAlaIleAspProAsnGIuProThrTyrcLysLysAsnGIuValSerTyrcGI 223
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 505 GACCTCCCGACATGCCCAAGAACCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 564
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 224 GIuMetIleGIuLysAspAsnGIuGIuCysProIleGIuTrpPheHisPheSerCysVal 243
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 565 GAGATGTGCGCTGCGACAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 624
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 244 SerLeuThrLysProLysGIuLysTrpTyrcLysProLysCysArgGIuAspAsnGIu 263
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 625 GGGCTCAATCATTAACCAAGGCGAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 684
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 264 LysThrMetAspLysSerThrGIuLysThrLysLysAspArg 277
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 685 AAGACATGTGCAAAAGCCCTGAGAAATTCAAAAAAGAGAG 726
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

# RESULT 12

AAV62285 standard; cDNA; 1902 BP.

AAV62285;

18-JAN-1999 (first entry)

Partial INGI partial cDNA sequence.

XX INGI gene; p33ING1; human; apoptosis; cell death; breast cancer;  
KW brain tumour; gene therapy; tumour suppressor; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 109..741

XX W09844102-A2.

XX 08-OCT-1998.

XX 26-MAR-1998; 98WO-CA00277.

XX 27-MAR-1997; 97US-0828158.

XX (UNITE-) UNIV TECHNOLOGIES INT INC.

XX Garkavtsev I, Helbing CC, Johnston RN, Riabowol K;

XX WPI; 1998-542700/46.

XX P-PSDB; AAW79674.

PT Modulating eukaryotic apoptosis by increasing p33ING1 activity -  
PT using p33ING1 derivatives, to induce apoptosis in cancer cells, and  
PT in the investigation of apoptotic pathways

PS Example 2; Fig 2; 66pp; English.

XX This is the nucleotide sequence of a human INGI (inhibitor of  
XX growth) partial cDNA clone that codes for a p33ING1 polypeptide  
XX (see AAW79674), a novel inhibitor of cell growth and a candidate  
XX tumour suppressor. INGI is a new gene that is expressed in normal  
XX mammary epithelial cells, but which is expressed only at lower  
XX levels in several cancerous mammary epithelial cell lines and is  
XX not expressed in many primary brain tumours. To isolate INGI, a  
XX subtractive hybridisation of breast cancer cell line cDNAs was  
XX performed with cDNA from normal mammary epithelial cells, and  
XX subcloned cDNAs were cloned into retrovirus vector pLNCX.  
XX Following passage through a packaging line, normal mouse mammary  
XX epithelial cells were infected, and infected cells were injected  
XX into nude mice. Putative transforming fragments from tumours were  
XX isolated by PCR (see AAV62290-91) and subcloned into LNCX. An INGI  
XX fragment was obtained and used to screen normal human fibroblast  
XX and HeLa cell cDNA libraries. 2 Clones were sequenced to obtain  
XX the partial INGI sequence. The complete cDNA sequence (see  
XX AAV62292) was obtained by RACE. A claimed method to potentiate  
XX apoptosis in a eukaryotic cell involves administering an active  
XX p33ING1 peptide or an oligonucleotide encoding such as a peptide.  
XX A claimed method for inhibiting apoptosis in a eukaryotic cell  
XX involves administering an antisense oligonucleotide. Also claimed  
XX are a method for determining the apoptotic characteristics of a  
XX eukaryotic cell, an assay for determining the level of p33ING1  
XX activity in a eukaryotic cell, and an isolated eukaryotic cell  
XX substantially free of p33ING1 biological activity. The invention  
XX discloses INGI derivatives or variants that may be used to induce  
XX apoptosis in eukaryotic cancer cells.

XX Sequence 1902 BP; 574 A; 390 C; 462 G; 476 T; 0 other;

# Alignment Scores:

Pred. No.:	Length:	1902
Score:	Matches:	136
Percent Similarity:	Conservative:	39
Best Local Similarity:	Mismatches:	45
Query Match:	Indels:	14
	Gaps:	2

US-09-513-365a-1 (1-280) x AAV62285 (1-1902)

OY 56 TyrcInGIuThrLeuLysGluIleAspAspValTyrcGluTyrcLysLysLysAsp 75



QY 144 GlnProGluArgSerArgArgProArgGlnArgGlnThrSerGluSerArgAspLeu 163  
 Db 490 AAGCCC-----AACAGCAAGCGCTCAAGCGGCGAGCCCAACAGAACCGTGAGAAC 543  
 QY 164 CysHisMetAlaAsnGlyIleGluAspCysAspAspGlnProProlGlyLysSer 183  
 Db 544 GCGTCCAGCAACCGACGACGACGAGCGCTCGGCGACACCCAGAGAGAGAGGCC 603  
 QY 184 LysSerAlaLysLysLysLysArgSerLysAlaLysGlnArgGluAlaSerProVal 203  
 Db 604 AAGACCTCCAGAGAGAGAGAGCGCTCCAGGCGGAGCGAGAGCGCTCCCTCC 663  
 QY 204 GluPheAlaIleAspProAsnGluProThrTyrcysLeuGlyAsnGlnValSerTyrcy 223  
 Db 664 GACCTCCCAACGACCCCAAGAACCCAGCTACTGTGTGTCCACACAGCTCTCTATGG 723  
 QY 224 GluMetIleGlyCysAspAsnGluGlnCysProIleGluTyrPheHisPheSerCysVal 243  
 Db 724 GAGATGATCGGCTCCGACACGACGAGCGCCCATCGAGTGTCTCCACTCTCGTGCCTG 783  
 QY 244 SerLeuThrTyrcysProLysGlyLysTyrcysProLysGlyAspAsnGlu 263  
 Db 784 GGGCTCAATCTTAACCCAGGCAAGTGTGTCTGTCTCCAGAGTCCGGGGGAGAGAG 843  
 QY 264 LysThrMetAspLysSerThrGluLysThrLysLysAspArg 277  
 Db 844 AAGACCATGAGCAAGCCCTGGAGAAATCCAAAAAGAGAG 885  
 RESULT 14  
 AAV62292 ID AAV62292 standard; cDNA; 2061 BP.  
 AC AAV62292;  
 XX 18-JAN-1999 (first entry)  
 DT Human INGI full-length cDNA sequence.  
 DE INGI gene; p3ING1; human; apoptosis; cell death; breast cancer;  
 KW brain tumour; gene therapy; tumour suppressor; ss.  
 XX Homo sapiens.  
 OS  
 FH Key Location/Qualifiers  
 FT CDS 16..900  
 FT /tag= a  
 XX MO5844102-A2.  
 PN 08-OCT-1998.  
 PD 26-MAR-1998; 98WO-CA00277.  
 PE 27-MAR-1997; 97US-0828158.  
 PR (UTTE-) UNIV TECHNOLOGIES INT INC.  
 PA Garkavtsev I, Helbing CC, Johnston RN, Riabowol K;  
 XX WPI: 1998-542700/46.  
 DR P-PSDB; AAW79675.  
 DR  
 PT Modulating eukaryotic apoptosis by increasing p3ING1 activity -  
 PT using p3ING1 derivatives, to induce apoptosis in cancer cells, and  
 PT in the investigation of apoptotic pathways  
 CC Claim 7; Fig 3; 66pp; English.  
 CC This is the nucleotide sequence of a human INGI (inhibitor of  
 CC growth) that codes for a p13ING1 polypeptide (see AAW79675), a novel  
 CC inhibitor of cell growth and a candidate tumour suppressor. INGI  
 CC is a new gene that is expressed in normal mammary epithelial cells,  
 CC but which is expressed only at lower levels in several cancers

CC mammary epithelial cell lines and is not expressed in many primary  
 CC brain tumours. To isolate INGI, a subtractive hybridisation of  
 CC breast cancer cell line cDNAs was performed with cDNA from normal  
 CC mammary epithelial cells, and subtracted cDNAs were cloned into  
 CC retrovirus vector pINCX. Following passage through a packaging  
 CC line, normal mouse mammary epithelial cells were infected, and  
 CC infected cells were injected into nude mice. Putative transforming  
 CC fragments from tumours were isolated by PCR (see AAV62290-91) and  
 CC subcloned into LNCX. An INGI fragment was obtained and used to  
 CC screen normal human fibroblast and HeLa cell cDNA libraries. 2  
 CC clones were sequenced to obtain a partial INGI sequence (see  
 CC AAV62285). The complete cDNA sequence was then obtained by RACE. A  
 CC claimed method to potentiate apoptosis in a eukaryotic cell  
 CC involves administering an active p3ING1 peptide or an  
 CC oligonucleotide encoding such as a peptide. A claimed method for  
 CC inhibiting apoptosis in a eukaryotic cell involves administering an  
 CC antisense oligonucleotide. Also claimed are a method for  
 CC determining the apoptotic characteristics of a eukaryotic cell, an  
 CC assay for determining the level of p3ING1 activity in a eukaryotic  
 CC cell, and an isolated eukaryotic cell substantially free of p3ING1  
 CC biological activity. The invention discloses INGI derivatives or  
 CC variants that may be used to induce apoptosis in eukaryotic cancer  
 CC cells.  
 CC XX  
 SQ Sequence 2061 BP; 602 A; 439 C; 515 G; 505 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 5,14e-60 Length: 2061  
 Score: 728.00 Matches: 136  
 Percent Similarity: 74.79% Conservative: 39  
 Best Local Similarity: 58.12% Mismatches: 45  
 Query Match: 49.16% Indels: 14  
 DB: 19 Gaps: 2  
 US-09-513-365a-1 (1-280) x AAV62292 (1-2061)  
 QY 56 TyrGlnGluThrLeuLysGluIleLeuAspValTyrGluLysGluLysAsp 75  
 Db 190 TGGAAACAGATCTTGAAGAGACTAGACAGTCTACGAGCGCTTACGTCGAGACAGAC 249  
 QY 76 LeuAsnGluLysLysArgLeuGlnGluLeuGlnArgAlaLeuIleAsnSerGlnGlu 95  
 Db 250 GGGGCGCAGAGAGCGGCGATCTGCATGTGTGACGCGCGCTGATCCGAGCAGAG 309  
 QY 96 LeuGlyAspGluLysIleGlnIleValThrGlnMetLeuGluLeuValGlnAspArgAla 115  
 Db 310 CTGGCGCAGCAGAGATCTCATGCTGAGAGTGTGAGTGTGTGAGAACCGCAGC 369  
 QY 116 ArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGlu----- 130  
 Db 370 CGGCAGGTGACAGCCAGCTGGAGCTGTTCAGAGCGGACGAGACTGGCGACAGCTG 429  
 QY 131 -----SerGluArgAlaSerAspLysAlaLysMetAspSer 143  
 Db 430 GCGAAGAGCGGCAAGTTGGCGGAGACAGGCCCAATGCGGTAGCGGAGTCTGAC 489  
 QY 144 GlnProGluArgSerArgArgProArgGlnArgGlnThrSerGluSerArgAspLeu 163  
 Db 490 AAGCCC-----AACAGCAAGCGCTCAAGCGGCGAGCCCAACAGAACCGTGAGAAC 543  
 QY 164 CysHisMetAlaAsnGlyIleGluAspCysAspAspGlnProProlGlyLysSer 183  
 Db 544 GCGTCCAGCAACCGACGACGACGAGCGCTCGGCGACACCCAGAGAGAGAGGCC 603  
 QY 184 LysSerAlaLysLysLysArgSerLysAlaLysGlnArgGluAlaSerProVal 203  
 Db 604 AAGACCTCCAGAGAGAGAGAGCGCTCCAGGCGGAGCGAGAGCGCTCCCTCC 663  
 QY 204 GluPheAlaIleAspProAsnGluProThrTyrcysLeuGlyAsnGlnValSerTyrcy 223  
 Db 664 GACCTCCCAACGACCCCAAGAACCCAGCTACTGTGTGTCCACACAGCTCTCTATGG 723  
 QY 224 GluMetIleGlyCysAspAsnGluGlnCysProIleGluTyrPheHisPheSerCysVal 243

```

Db      724 GAGATGATCGGCGCACAACGAGAGTCCCATGAGTGGTCCACTTCTGTCGTG 783
QY      244 SerLeuThrTyrLysProLysGlyLysTyrPtyrCysProLysCysArgGlyAspAsnGlu 263
Db      784 GGCGCTAAATCAATAAACCCAAAGGCAAGTGTACTGTCTCCAAAGTCCGGGGGAGAACGAG 843
QY      264 LysThrMetAspLysSerThrGluLysThrLysLysAspArg 277
Db      844 AAGACCATGTGACAAAGCCCTGGAGAAATCCAAAAGAGAGG 885

RESULT 15
AAx28688
ID      AAx28688 standard; cDNA; 2061 BP.
XX
AC      AAx28688;
XX
DT      26-AUG-1999 (first entry)
XX
DE      Nucleotide sequence of the human p33-ING1 gene.
XX
KW      Human; p33-ING1 protein; growth regulation; apoptosis; DNA damage;
KW      inhibition; anchorage independent growth; cytotoxic drug;
KW      transcriptional activation; cancer; immortal cell line; ss.
XX
OS      Homo sapiens.
XX
FH      Key
FT      CDS
FT      16..900
FT      /tag= a
FT      /product= "human p33-ING1 protein"
XX
XX      WC0916790-A1.
XX      08-APR-1999.
XX
PF      24-SEP-1998; 98WO-US18179.
XX
PR      14-JAN-1998; 98US-0006783.
PR      26-SEP-1997; 97US-0060138.
XX
PA      (UNIT ) UNITV ILLINOIS BOARD OF TRUSTEES.
PA      (UNIT-) UNITV TECHNOLOGIES INT INC.
XX
PI      Garkavtsev I, Gudkov A, Riabowol K;
XX
DR      MPI; 1999-263685/22.
XX
PT      P-PSDB; AAY03244.
XX
XX      Use of p33-ING1 peptides
XX
PS      Disclosure; Fig 1; 64pp; English.
XX
XX      This is the nucleotide sequence encoding the human p33-ING1 protein,
XX      which can be used to modulate the activity of, isolate or detect p53.
XX      Expression of the ING1 and p53 genes in a mammalian cell results in
XX      normal growth regulation anchorage-dependent growth and apoptosis as
XX      a response to irreversible DNA damage and other cellular insult.
XX      Inhibition of expression of either gene results in a loss of cellular
XX      growth control, anchorage independent growth, inhibition of apoptosis
XX      and resistance to radiation and cytotoxic drugs. The p33-ING1 is a
XX      component of the p53 signalling pathway that cooperates with p53 in
XX      negative regulation of cell proliferation by modulating p53 signalling
XX      transcriptional activation. Biological function of p53 signalling
XX      pathway can therefore be regulated (both enhanced or suppressed) by
XX      modulating p33-ING1 activity. The modulation of p33-ING1 activity can
XX      be used for the stimulation or restoration of the p53 pathway in
XX      anti cancer therapy or for the suppression of the p53 pathway to
XX      defend sensitive tissues from genotoxic stress or for the generation
XX      of immortal cell lines.
XX
SQ      Sequence 2061 BP; 600 A; 437 C; 518 G; 506 T; 0 other;

```

```

Alignment Scores:
Pred. No.: 5,14e-60 Length: 2061
Score: 728.00 Matches: 136
Percent Similarity: 74.79% Conservative: 39
Best Local Similarity: 58.12% Mismatches: 45
Query Match: 49.16% Indels: 14
DB: 20 Gaps: 2

US-09-513-365a-1 (1-280) x AAx28688 (1-2061)
QY      56 TyrGlnGluThrLysGluLysGluLeuLeuAspValTyrGluLysTyrLysGluAspAsp 75
Db      190 TGGAAACAGATCCTGAAGGAGCTAGACAGAGTGTACGAGCGCTTCAGTCCGAGACAGAC 249
QY      76 LeuAsnGlnLysLysArgLeuGlnGlnLeuLeuGlnArgAlaLeuLeuLeuSerGlnGlu 95
Db      250 GGGGCGCAGAAAGGGGGGAGTGTGACACTGTGTGACGCGCGCTGTATCCGACCGACAGAG 309
QY      96 LeuGlyAspGluLysIleGlnIleValThrGlnMetLeuGlnLeuValGluAsnArgAla 115
Db      310 CTGGGGGAGACAGAAATCCAGATCGTAGAGCCAGATGTGTGAGTGTGGAGAACCCAGC 369
QY      116 ArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGlu----- 130
Db      370 CGGCGAGGTGACACGACAGTGTGAGCTGTTCGAGCGCAGCAGCAGTGGGCGACAGAGTG 429
QY      131 -----SerGluArgAlaSerAspLysAlaLysMetAspSer 143
Db      430 GGCACACAGCGGCAAGTGTGGCGGAGAGGCCCAATGGCGATCGGTAGCGCATCTGAC 489
QY      144 GlnProGluArgSerArgArgProArgArgGlnArgThrSerGlnSerArgAspLeu 163
Db      490 AAGCCC-----AACAGCAAGCGCTCAGCGGCGAGCGCAACAGCAAGCAAGCGTGAGAAC 543
QY      164 CysHisMetAlaAsnGlyIleGluAspCysAspAspGlnProProLysGluLysSer 183
Db      544 GCGTCCAGCAACCCAGCAGCAGAGCGGCGCTCGGCGACACCCCAAGGAGGAAGAGGCC 603
QY      184 LysSerAlaLysLysLysArgSerLysAlaLysGlnGlnArgGlnAlaSerProVal 203
Db      604 AAGACCTCCCAAGAGAAAGCGCTCCAGGCCAGAGCGGAGAGGCGTCCCTGCC 663
QY      204 GluPheAlaIleAspProAsnGlnProThrTyrCysLeuGlnValSerTyrGly 223
Db      664 GACCTCCCATCGACCCCAAGCAAGCCAGTGTCTGTGCAACAGGCTCTCTATWGG 723
QY      224 GluMetIleGlyCysAspAsnGlnGlnCysProIleGluTrpPheHisPheSerCysVal 243
Db      724 GAGATGATCGCTGCGACAAAGAGTGGCCCATGAGTGTTCACACTTCTGTCGTG 783
QY      244 SerLeuThrTyrLysProLysGlyLysTyrPtyrCysProLysCysArgGlyAspAsnGlu 263
Db      784 GGCGCTAAATCAATAAACCCAAAGGCAAGTGTACTGTCTCCAAAGTCCGGGGGAGAACGAG 843
QY      264 LysThrMetAspLysSerThrGluLysThrLysLysAspArg 277
Db      844 AAGACCATGTGACAAAGCCCTGGAGAAATCCAAAAGAGAGG 885

```

Search completed: June 20, 2003, 02:49:55  
Job time : 250 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 20, 2003, 02:14:54 ; Search time 2153 Seconds  
(without alignments)  
3784.852 Million cell updates/sec

Title: US-09-513-365A-1  
Perfect score: 1481  
Sequence: 1 MLGQQQQQLYSSAALLTGER.....DNEKTMKSTKTKKDRRSR 280

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2.1/USPTO\_spool/US09513365/runat\_16062003\_124743\_16442/app\_query.fasta\_1.455  
-DB=GenEmbl -QFMT=fastcap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09513365.ecgn.1.1.3745@runat\_16062003\_124743\_16442 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl :  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*

29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.rod.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Length	DB ID	Description
1	1481	100.0	1078	9	AB012853 Homo sapi
2	1481	100.0	1080	6	AX211560 Sequence
3	1481	100.0	1080	9	AF053537 Homo sapi
4	1481	100.0	1141	9	BC030128 Homo sapi
5	1478	99.8	843	9	AJ006851 Homo sapi
6	1399.5	94.5	1001	10	AF078834 Mus muscu
7	1197	80.8	974	9	HSING2S2
8	1197	80.8	189972	2	AC107214 Homo sapi
9	1152	77.8	168245	2	AC128348 Rattus no
10	1041	70.3	1094	5	AY014017 Xenopus l
11	1008	68.1	42029	2	AC114762 Homo sapi
12	880.5	59.5	840	10	AF149820 Mus muscu
13	874.5	59.0	1584	10	BC016573 Mus muscu
14	874.5	59.0	2817	10	AF177757 Mus muscu
15	853.5	57.6	840	9	AF078835 Homo sapi
16	853.5	57.6	845	9	AF310392 Homo sapi
17	853.5	57.6	873	6	AR171884 Sequence
18	853.5	57.6	911	9	AF044076 Homo sapi
19	853.5	57.6	1533	6	AX367042 Sequence
20	853.5	57.6	1533	9	AF149721 Homo sapi
21	853.5	57.6	2444	9	AB024401 Homo sapi
22	853.5	57.6	2897	9	AF181850 Homo sapi
23	750	50.6	1944	10	AF177756 Mus muscu
24	750	50.6	1959	10	AF177755 Mus muscu
25	749	50.6	235079	2	AC124475 Mus muscu
26	748	50.5	166219	2	AC120896 Rattus no
27	730	49.3	1074	9	AF149722 Homo sapi
28	728	49.2	1143	6	AX367043 Sequence
29	728	49.2	1189	9	AF149723 Homo sapi
30	728	49.2	1902	6	AR079046 Sequence
31	728	49.2	1902	6	AR087457 Sequence
32	728	49.2	1902	6	AR110646 Sequence
33	728	49.2	1902	6	AR154563 Sequence
34	728	49.2	1980	9	AB024404 Homo sapi
35	728	49.2	2061	6	AB087464 Sequence
36	728	49.2	2061	6	AR110653 Sequence
37	728	49.2	2061	6	AR154570 Sequence
38	728	49.2	2061	9	AF001954 Homo sapi
39	728	49.2	2096	9	HSP333ING2
40	728	49.2	2296	9	BC018348 Homo sapi
41	728	49.2	2306	9	AB031269 Homo sapi
42	728	49.2	2552	9	HSING3
43	728	49.2	2886	9	AF181849 Homo sapi
44	728	49.2	2891	9	AB024402 Homo sapi
45	728	49.2	2925	9	AB037387 Homo sapi

ALIGNMENTS

RESULT 1

```

AB012853      1078 bp      mRNA      linear      PRI 18-MAR-1999
LOCUS          Homo sapiens INGL1 mRNA for INGL1p, complete cds.
DEFINITION
ACCESSION      AB012853
VERSION        AB012853.1 GI:4115554
KEYWORDS       INGL1p.
SOURCE          Homo sapiens fetal tissue_lib:fetal-brain cDNA to mRNA.
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (sites)
AUTHORS        Shimada,Y., Saito,A., Suzuki,M., Takahashi,E. and Horie,M.
TITLE          Cloning of a novel gene (INGL1) homologous to INGL1, a candidate
               tumor suppressor
JOURNAL        Cytogenet. Cell Genet. 83 (3-4), 232-235 (1998)
MEDLINE        99172097
REFERENCE      2 (bases 1 to 1078)
AUTHORS        Shimada,Y., Saito,A. and Horie,M.
TITLE          Direct Submission
JOURNAL        Submitted (07-APR-1998) Yoshikazu Shimada, Otsuka Pharmaceutical
               Co. Ltd., Otsuka GEN Research Institute, Kagasuno, Kawauchi-cho,
               Tokushima, Tokushima 771-0192, Japan
               (E-mail:shim@otsuka.genome.ad.jp, Tel:81-886-65-2888(ex.2476),
               Fax:81-886-37-1035)
COMMENT        Sequence updated (17-Apr-1998).
FEATURES       Location/Qualifiers
               1..1078
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="4"
                /map="4q35.1"
                /dev_stage="fetal"
                /tissue_lib="fetal-brain"
               1..1078
                /gene="INGL1"
                /gene="INGL"
                /codon_start=1
                /product="INGLp"
                /protein_id="BAA36419.1"
                /db_xref="GI:4115555"
                /translation="MLGQQQOLYSNAALLTGERSLLTCYVDLYECVESLPHDMQR
                NYSVLRELNKKYQELIKEIDDDYKYEKKEDDLNKKRLQQLLQRLINSGELGDEKIQ
                IVTOMLELVENKARQELHSCQFDDPAESERASDKAMDSQSPERSRRPRORTSES
                RDLCHWANGIEDDDPPPKSKSAKKRKAQOERASPFVEAIDNPRTYCLCN
                QVSYGMICGDNQEQCEIWFHFSCVSLTYKPKGWYCPKCRGDNKXTMDKSTKTKKD
                RRSR"
BASE COUNT    356 a 206 c 280 g 236 t
ORIGIN
Alignment Scores:
Pred. No.:    6.08e-113      Length:    1078
Score:        1481.00      Matches:    280
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:  100.00%      Indels:    0
DB:           9      Gaps:    0

US-09-513-365A-1 (1-280) x AB012853 (1-1078)
QY      1 MetLeuGlnGlnGlnGlnGlnLeuTyrSerSerAlaAlaLeuLeuThrGlyGluArg 20
Db      92 ATGTTAGGCAGCAGCAGCAGCAACTGTACTCGTCGGCGCGCTCTCGACGGGAGCGG 151
QY      21 SerArgLeuLeuThrCysTyrValGlnAspTyrLeuGluCysValGlnSerLeuProHis 40
Db      152 AGCCGCGCTGCTACCTGCTACGTCGAGGACTACTCTTGAGTCGCGGAGTCGCGCCAC 211
QY      41 AspMetGlnArgAsnValSerValLeuArgGluLeuAspAsnLysTyrGlnGluThrLeu 60
Db      212 GACATCAGAGGAACGCTGCTGCTGCGAGAGCTGCACAAACAAATATCAAGAAAGCTTA 271
QY      61 LysGluLeuAspAspValTyrGluLysTyrLysLysGluAspAspLeuAsnGlnLysLys 80

```

---

```

|||||
272 AAGGAAATGATGATGCTCTACGAAAAATATAAGAAAGACGATGATTTAAACCAAGAAA 331
QY      81 ArgLeuGlnGlnLeuLeuGlnArgAlaLeuIleAsnSerGlnGluLeuGlyAspGluLys 100
Db      332 CGTCTACAGCAGCTTCTCCAGAGGACACTAATTAATAGTCAAGATTTGGGAGATGAAAA 391
QY      101 IleGlnIleValThrGlnMetLeuLeuValGluAsnArgAlaArgGlnMetGluLeu 120
Db      392 ATACAGATTGTTACACAAATGCTCGAATGTGTGGAAATCGGCAAGACAAATGGAGTTA 451
QY      121 HisSerGlnCysPheGlnAspProAlaGluSerGluArgAlaSerAspLysAlaLysMet 140
Db      452 CACTCACAGTGTTCCTCAAGATCCTGCTGAAAGTGAACGAGCCTCAGATAAAGCAAGATG 511
QY      141 AspSerSerGlnProGluArgSerSerArgArgProArgArgGlnArgThrSerGluSer 160
Db      512 GATTTCAGCCCAACGAGAAAGATCTTCAAGAAAGCCCGCAGGAGGAGCAGTGAAGC 571
QY      161 ArgAspLeuCysHisMetalAlaAsnGlyIleGluAspCysAspAspGlnProProLysGlu 180
Db      572 CGTGATTATGTCACATGGCAATGGGATTGAAGACTGTGATGATCAGCCACCTAAAGAA 631
QY      181 LysLysSerLysSerAlaLysLysLysLysLysLysLysLysLysLysGlnGluArgGluAla 200
Db      632 AAGAAATCCCAAGTCAGCAAAAGAAAGAAACGCTCCAAAGGCCAAGCAGGAGGAAAGAGCT 691
QY      201 SerProValGluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnVal 220
Db      692 TCACCTGTTGAGTTTGCAATAGATGCTTAATGAACCTACATACACTGCTTATCAACCAAGT 751
QY      221 SerTyrGlyGluMetIleGlyCysAspAsnGluGlnCysProIleGluTrpPheHisPhe 240
Db      752 TCATTATGGGAGATGATAGATGTGACATGAACAGTGTCCATTGAATGTTTCACTTT 811
QY      241 SerCysValSerLeuThrTyrLysProLysGlyLysTyrTrpTyrCysProLysCysArgGly 260
Db      812 TCATGTGTTTCACTTACCTTACCTATAACCAAGGGGAAATGGTATTGCCCAAAAGTCAGGG 871
QY      261 AspAsnGluLysThrMetAspLysSerThrGluLysThrLysLysLysAspArgSerArg 280
Db      872 GATATGAGAAAACAATGGCAAAAGTACTGAAAGACAAAAGAGATAGATAGATCGAGG 931

RESULT 2
AX211560      1080 bp      DNA      linear      PAT 06-SEP-2001
LOCUS          Sequence 7 from Patent WO0159114.
DEFINITION
ACCESSION      AX211560
VERSION        AX211560.1 GI:15523814
KEYWORDS       human.
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 1080)
AUTHORS        Harris,C.C. and Nagashima,M.
TITLE          Tumour suppressor gene, p47ing3
JOURNAL        Patent: WO 0159114-A 7 16-AUG-2001;
               THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)
FEATURES       Location/Qualifiers
               source
                1..1080
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /note="p33ING2"
BASE COUNT    375 a 196 c 271 g 238 t
ORIGIN
Alignment Scores:
Pred. No.:    6.09e-113      Length:    1080
Score:        1481.00      Matches:    280
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:  100.00%      Indels:    0

```







```

QY 181 LysLysSerLysSerAlaLysLysLysLysArgSerLysAlaLysGlnGluArgGluAla 200
Db 676 AAGAAATCCAGTCAGCAAGAAAGAAAGAAAGCGCTCCCAAGCCCAAGCAAGAAAGGAAGCT 735

QY 201 SerProValGluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnVal 220
Db 736 TCACCTGTTGAGTTGCAATAGATTCCTAATGAACTACATACCTGCTTATGCAACCAAGTG 795

QY 221 SerTyrGlyGluMetIleGlyCysAspAsnGluGlnCysProIleGluTTPPheHisPhe 240
Db 796 TCTATTGGGGAGATGATAGATGTGACATGACAGTGTCCCAATGATGTTTCACTTT 855

QY 241 SerCysValSerLeuThrTyrLysProLysGlyLysTyrCysProLysCysArgGly 260
Db 856 TCATGTGTTTCACTTACCTATATAACCAAGGGAATGTTATGCCCAAGTGCAGGGGA 915

QY 261 AspAsnGluLysThrMetAspLysSerThrGluLysThrLysLysAspArgSerArg 280
Db 916 GATAATGAGAAACAATGGACAAAGTACTGAAAGACAAACAAAGGATAGATCGAGG 975

RESULT 5
HSAJ6851      843 bp      mRNA      linear      PRI 06-JAN-2001
LOCUS
DEFINITION Homo sapiens mRNA for p32 protein.
ACCESSION AJ006851
VERSION AJ006851.1 GI:12053587
KEYWORDS ing2 gene; p32 protein.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 843)
AUTHORS Cal.S., Freije,J.M. and Lopez-Otin,C.
TITLE ING2,a new possible gene supressor tumor
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 843)
AUTHORS Lopez-Otin,C.
TITLE Direct Submission
JOURNAL Submitted (08-JUN-1998) Lopez-Otin C., Bioquimica y Biologia Molecular, Universidad de Oviedo. Facultad de Medicina., C/ Julian Claveria s/n., 33006, SPAIN
FEATURES
source location/Qualifiers
1..843
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="breast tumor"
1..843
/gene="ing2"
1..843
/gene="ing2"
/codon_start=1
/product="p32 protein"
/protein_id="CAC20567.1"
/db_xref="GI:12053588"
/translation="MLGQQQQLYSSAALTGERTRLLTCYVDYLCEVSLPHDMQR
NVSLRELDNKGQETLKEIDVYKKEDDLKQKRLQQLLRALINSQELDEKIQ
IVTQMLEVNRAROMELHSCQFDPAESRSDKAKMSSQPSRRRPRORTSES
RDLCHWANGIEDCDOPPEKKSKSAKKRKAOKEREASPVFEAIDPNEPTCLCN
QVSYGEMIGCDNQCEPTEWFHFCVSLTYPKGKNYCPKCRGDNEXTMDKSTKTKD
RRSR"
BASE COUNT 298 a 163 c 212 g 170 t
ORIGIN
Alignment Scores:
Pred. No.: 7.87e-113 Length: 843
Score: 1478.00 Matches: 279
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.64% Mismatches: 0
Query Match: 99.80% Indels: 0
DB: 9 Gaps: 0
US-09-513-365A-1 (1-280) x HSAJ6851 (1-843)

```

```

QY 1 MetLeuGlyGlnGlnGlnGlnLeuTyrSerSerAlaAlaLeuLeuThrGlyGluArg 20
Db 1 ATCTAGGGCAGCAGCAGCAGCAACCTGCTACTCGCTCGCTCGCTCGCTCGCTCGCG 60

QY 21 SerArgLeuLeuThrCysTyrValGlnAspTyrLeuGluCysValGluSerLeuProHis 40
Db 61 ACCGGCTGCTCAGCTGCTAGTGCAGGAGCTACCTTGAGTGCCTGGAGTGCCTGCCAC 120

QY 41 AspMetGlnArgAsnValSerValLeuArgGluLeuAspAsnLysTyrGlnGluThrLeu 60
Db 121 GACATGCAGAGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180

QY 61 LysGluIleAspAspValTyrGluLysLysLysGluAspAspLeuAsnGlnLysLys 80
Db 181 AAGGAAATTCATGATGCTACGAAATAATAAGAAAGCAAGATGATTTAAACCAAGCAAA 240

QY 81 ArgLeuGlnGlnLeuLeuGlnArgAlaLeuIleAsnSerGlnGluLeuGlyAspGluLys 100
Db 241 CGTCTACAGCAGCTTCTCCAGAGAGCACTAATTAATAGTCAAGAATTTGGAGATGAAA 300

QY 101 IleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAlaArgGlnMetGluLeu 120
Db 301 ATACAGATTTGTTACACAAATGCTCGAATTTGTTGGAATAATCGGCGACACAAATGGAGTTA 360

QY 121 HisSerGlnCysPheGlnAspProAlaGluSerGluArgAlaSerAspLysAlaLysMet 140
Db 361 CACTCACAGTGTTCCTCAAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420

QY 141 AspSerSerGlnProGluArgSerSerArgArgProArgArgGlnArgThrSerGluSer 160
Db 421 GATTCACGCAACCCAGAAAGATCTTCAAGAAAGACCCCGCAGGCGACGACGACGTTGAAGC 480

QY 161 ArgAspLeuCysHisMetAlaAsnGlyIleGluAspCysAspAspGlnProProLysGlu 180
Db 481 CGTGATTTATGTCACATGGCAATGGGATTTGAACACTGTGATGATGATGATGATGATGAT 540

QY 181 LysLysSerLysSerAlaLysLysLysLysArgSerLysAlaLysLysGlnGluArgGluAla 200
Db 541 AAGAAATCCAAAGTCAGCAAGAAAGAAAGAAAGCGCTCCCAAGGCGCAAGCAAGGAAAGGAGCT 600

QY 201 SerProValGluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnVal 220
Db 601 TCACCTGTTGAGTTTGAATAGATCTCAATGAAGCTACATCTCTCTTATGCAACCAAGTG 660

QY 221 SerTyrGlyGluMetIleGlyCysAspAsnGlnCysProIleGluTTPPheHisPhe 240
Db 661 TCTATGGGAGATGATAGATGTGACAAATGAACAGTGTCCCAATGATGATGATGATGATGAT 720

QY 241 SerCysValSerLeuThrTyrLysProLysGlyLysTyrCysProLysCysArgGly 260
Db 721 TCATGTGTTTCACTTACCTATATAACCAAGGGAATGTTATGCTCCCAAGTGCAGGGGA 780

QY 261 AspAsnGluLysThrMetAspLysSerThrGluLysThrLysLysAspArgSerArg 280
Db 781 GATAATGAGAAACAATGGACAAAGTACTGAAAGACAAACAAAGGATAGAGATCGAGG 840

RESULT 6
AF078834      1001 bp      mRNA      linear      ROD 09-SEP-2000
LOCUS
DEFINITION Mus musculus p33ING2 (Ing2) mRNA, complete cds.
ACCESSION AF078834
VERSION AF078834.1 GI:10039542
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1001)
AUTHORS Nagashima,M., Hagiwara,K., Hancock,A.R. and Harris,C.C.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-1998) Laboratory of Human Carcinogenesis, National Cancer Institute, National Institutes of Health, 37 Convent Dr., Bldg. 37, Rm. 2C01, Bethesda, MD 20892, USA

```

[illegible]

MO 63108, USA  
3 (bases 1 to 189972)  
Waterston, R.H.

Direct submission  
Submitted (10-APR-2002  
University School of M  
MO 63108, USA

On Apr 10, 2002 this S  
----- Ge  
Center: Washington Uni

Center code: WUGSC  
Web site: <http://genome>  
Contact: [submissions@w](mailto:submissions@w)  
----- proje

```

center project name: H
-----
----- Summary
Sequencing vector: M13

```

Sequencing vector: pUC19  
Chemistry: Dye-primer  
Chemistry: Dye-termina  
Assembly program: Phra

Consensus quality:	189
Consensus quality:	190
Consensus quality:	191

Insert size: 195974; s  
Quality coverage: 11.3  
Quality coverage: 10.9

\* NOTE: This is a 'wor  
\* consists of 8 contig  
\* is not known and the

\* runs of N, but the e  
\* This record will be  
\* as soon as it is ava

*	1	1252: C
*	1253	1352: g
*	1353	3199: C

*	3300	4420:	C
*	4421	4520:	g
*	4521	5658:	C

*	5759	6866: C
*	6867	6966: g
*	6967	37117: C

37218	101402: g
37218	101502: g
101403	189972: c
101503	

```
1. .189972
/organism="Homo sapiens"
/db_xref="taxid:9606"

```

```
feature
1. .1252
/clone="RP11-
/chromosome=
/pote="assemb
```

```

133. :3199
/note="assemb
3300. :4420
/note="assemb

```

ing Center, Washington  
Forest Park Parkway, St. Louis,

ing Center. Washington

replaced gl:18677601.

Sequencing Center  
index.shtml[illegible]

1

0% of reads  
0319  
ast Q40 .  
4.28

ast Q20

sum-of-contigs  
-----  
quence. It currently

s sequence record is  
are represented as  
the gaps are unknown.

Accession number will  
be in length

length  
p in length  
length  
p in length

length  
p in length  
length  
p in length

length  
top in length  
length  
top in length

length  
top in length.

misc\_feature 4521..5658  
/note="assembly\_name:Contig56"  
misc\_feature 5759..6866  
/note="assembly\_name:Contig58"  
misc\_feature 6967..73117  
/note="assembly\_name:Contig65"  
misc\_feature 73218..101402  
/note="assembly\_name:Contig66"  
clone\_end:SP6  
vector\_side:left  
101503..189972  
/note="assembly\_name:Contig67"  
clone\_end:T7  
vector\_side:left  
BASE COUNT 52803 a 40308 c 41211 g 54950 t 700 others  
ORIGIN

Alignment Scores:  
Pred. No.: 8,95e-87 Length: 189972  
Score: 1197.00 Matches: 231  
Percent Similarity: 93.98% Conservative: 3  
Best Local Similarity: 92.77% Mismatches: 9  
Query Match: 80.82% Indels: 6  
DB: 2 Gaps: 2

US-09-513-365A-1 (1-280) x AC107214 (1-189972)

QY 32 LeuGluAspAspLeuGluSerLeuProHisAspMetGlnArgAsnValSerValLeuArgGlu 51  
||||| : : : : : ||||| : : : : :  
146651 TTGGAATGTTGCTGCTACACATGAT-----AACGTTCTCTTTTCTTTTC 146701  
QY 52 LeuAspAsnLysTyrglnGluThrLeuLysGluLeuAspValTyrglnLysTyrgln 71  
||| : : : : : ||||| : : : : :  
146702 CTT-----TTTTAGAACGTTAAAGGAATGATGATGCTACGAAATAAAG 146752  
QY 72 LysGluAspAspLeuAsnGlnLysLysArgLeuGlnGlnLeuLeuGlnArgAlaLeu 91  
||||| : : : : : ||||| : : : : :  
146753 AAAGAAGATGATTTAAACCAAGAAACGCTACACAGCTTCTCCAGAGACACTAAT 146812  
QY 92 AsnSerGlnGluLeuGluGlyAspGluLysIleGlnIleValThrGlnMetLeuGluLeuVal 111  
||||| : : : : : ||||| : : : : :  
146813 ATAGTCAGAAATGGAGATGAAATAATACAGATGTTACACAAATGCTGAAATGGTG 146872  
QY 112 GluAsnArgAlaArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGluSer 131  
||||| : : : : : ||||| : : : : :  
146873 GAAATCGGCGAAGACAAATGGAGTTACACTCACAGTGTTCAGATCTCTGTAAGT 146932  
QY 132 GluArgAlaSerAspLysAlaLysMetAspSerSerGlnProGluArgSerArgArg 151  
||||| : : : : : ||||| : : : : :  
146933 GAACGAGCTCAGATAAAGCAAGATGGATTCAGCCCAACCAAGAAAGATCTTCAAGAAGA 146992  
QY 152 ProArgGlnArgThrSerGluSerArgAspLeuCysHisMetAlaAsnGlyIleGlu 171  
||||| : : : : : ||||| : : : : :  
146993 CCCCAGGCGGACGACAGTGAACCGCTGATTTATGTCACATGCAATGGGATTTGAA 147052  
QY 172 AspCysAspAspGlnProProLysGluLysLysSerLysSerAlaLysLysLysArg 191  
||||| : : : : : ||||| : : : : :  
147053 GACTGTGATGATCAGCACCTAAAGAAAGAAATCCCAAGTCAGCAAGAAAGAAACGC 147112  
QY 192 SerLysAlaLysGlnGluArgGluAlaSerProValGluPheAlaIleAspProAsnGlu 211  
||||| : : : : : ||||| : : : : :  
147113 TCCAGGCCAACGAGAAAGGAAGCTTCCACTGTTGAGTTTGCATATAGATCCTAATGAA 147172  
QY 212 ProThrTyrcysLeuCysAsnGlnValSerTyrglnMetIleGlyCysAspAsnGlu 231  
||||| : : : : : ||||| : : : : :  
147173 CCTACATAGTCTTATGCAACCAAGTCTTTATGGGAGATGATAGATGACAAAGAA 147232  
QY 232 GlnCysProIleGluTrpPheHisPheSerCysValSerLeuThrTyrglnLysGly 251  
||||| : : : : : ||||| : : : : :  
147233 CAGTGTCCAATGAATGGTTTCACTTTTCATGTTTCACTTACCTATAAACCAGAGGG 147292  
QY 252 LysTrpTyrcysProLysCysArgGlyAspAsnGluLysThrMetAspLysSerThrGlu 271  
||||| : : : : : ||||| : : : : :

Db 147293 AAATGCTATTGCCCAAAAGTCAGGGGAGATATGAGAAAACAATGACAAAAGTACTGAA 147352  
QY 272 LysThrLysLysAspArgArgSerArg 280  
||||| : : : : : ||||| : : : : :  
Db 147353 AAGACAAAAGAGTAGAAGATCGAGG 147379  
RESULT 9  
AC128348 168245 bp DNA linear HTG 19-JUL-2002  
LOCUS Rattus norvegicus clone CH230-129P3, \*\*\* SEQUENCING IN PROGRESS  
DEFINITION \*\*\* 50 unordered pieces.  
AC128348  
VERSION AC128348.1 GI:21908993  
HTG: HTGS\_PHASE1.  
KEYWORDS Rattus norvegicus.  
SOURCE Rattus norvegicus  
ORGANISM Rattus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
REFERENCE 1 (bases 1 to 168245)  
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
Alsbrooks,S.L., Amaratunga,H.C., Are,J.R., Ayele,M., Banks,T.,  
Barbaria,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,  
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,  
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
Devilla,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hoggues,M., Holloway,C., Hollins,B.,  
Homi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
Kratochvil,J., Kureshi,A., Landry,N., Leal,B., Lewis,L., Lewis,L.,  
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louseged,H.,  
Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,  
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okunolu,G.,  
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G.,  
Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,  
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Tusmanik,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
Williams,G., Williamson,A., Wleczky,R., Wooden,S., Worley,K.,  
Wu,C., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G. and Gibbs,R.  
Direct Submission  
Unpublished  
2 (bases 1 to 168245)  
Worley K.C.  
Direct Submission  
Submitted (19-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu

TITLE  
JOURNAL  
Unpublished  
AUTHORS  
Worley K.C.  
TITLE  
JOURNAL  
COMMENT

----- Project Information  
Center project name: GZCX  
Center clone name: CH230-129p3  
----- Summary Statistics  
Sequencing vector: Plasmid;  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 115858 bases at least Q40  
Consensus quality: 122576 bases at least Q30  
Consensus quality: 127679 bases at least Q20  
-----

\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 50 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1  
1084: contig of 1084 bp in length  
1085  
1184: gap of unknown length  
1185  
2200: contig of 1016 bp in length  
2201  
3302: contig of 1002 bp in length  
3303  
3402: gap of unknown length  
3403  
4508: contig of 1106 bp in length  
4509  
5989: contig of 1381 bp in length  
5990  
6089: gap of unknown length  
6090  
7785: contig of 1696 bp in length  
7786  
9044: contig of 1159 bp in length  
9045  
9145  
10325: contig of 1181 bp in length  
10326  
10425: gap of unknown length  
10426  
11588: contig of 1163 bp in length  
11589  
11688: gap of unknown length  
11689  
13187: contig of 1499 bp in length  
13188  
13287: gap of unknown length  
13288  
14806: contig of 1519 bp in length  
14807  
14906: gap of unknown length  
14907  
15982: contig of 1076 bp in length  
15983  
16082: gap of unknown length  
16083  
18040: contig of 1958 bp in length  
18041  
18140: gap of unknown length  
18141  
18750: contig of 1610 bp in length  
18751  
19850: gap of unknown length  
19851  
21035: contig of 1185 bp in length  
21036  
21135: gap of unknown length  
21136  
22589: contig of 1454 bp in length  
22590  
22689: gap of unknown length  
22690  
23848: contig of 1159 bp in length  
23849  
25639: gap of unknown length  
25640  
25739: gap of unknown length  
25740  
28181: contig of 2442 bp in length  
28182  
28281: gap of unknown length  
28282  
29975: gap of unknown length  
29976  
32032: contig of 2057 bp in length  
32033  
33132: gap of unknown length  
33133  
33717: contig of 1585 bp in length  
33718  
33817: gap of unknown length  
33818  
35711: contig of 1894 bp in length  
35712  
35811: gap of unknown length  
35812  
37778: contig of 1967 bp in length  
37779  
37878: gap of unknown length  
37879  
40213: contig of 2335 bp in length  
40214  
40313: gap of unknown length  
40314  
42821: contig of 2514 bp in length  
42822  
42927: gap of unknown length

\* 42928  
42929: contig of 2343 bp in length  
43070: gap of unknown length  
43071  
48107: contig of 2737 bp in length  
48108  
48207: gap of unknown length  
48208  
51185  
51185: gap of unknown length  
51185  
53393: contig of 2109 bp in length  
53394  
53493: gap of unknown length  
53494  
56791: contig of 3298 bp in length  
56792  
56891: gap of unknown length  
56892  
62150: contig of 5259 bp in length  
62151  
62250: gap of unknown length  
62251  
65330: contig of 3080 bp in length  
65331  
65430: gap of unknown length  
65431  
69898: contig of 4468 bp in length  
69899  
69998: gap of unknown length  
69999  
74110: contig of 4112 bp in length  
74111  
76637: contig of 2427 bp in length  
76638  
76737: gap of unknown length  
76738  
82625: contig of 5888 bp in length  
82626  
82725: gap of unknown length  
82726  
85396: contig of 2671 bp in length  
85397  
85496: gap of unknown length  
85497  
89571: contig of 4075 bp in length  
89572  
89671: gap of unknown length  
89672  
94810: contig of 5139 bp in length  
94811  
94910: gap of unknown length  
94911  
98354: contig of 3444 bp in length  
98355  
98454: gap of unknown length  
98455  
103316: contig of 4862 bp in length  
103317  
103416: gap of unknown length  
103417  
108022: contig of 4606 bp in length  
108023  
108122: gap of unknown length  
108123  
114307: contig of 6185 bp in length  
114308  
114407: gap of unknown length  
114408  
121007: contig of 6600 bp in length  
121008  
121107: gap of unknown length  
121108  
127326: contig of 6219 bp in length  
127327  
127426: gap of unknown length  
127427  
136654: contig of 9228 bp in length  
136655  
136754: gap of unknown length  
136755  
148193: contig of 11439 bp in length  
148194  
148293: gap of unknown length  
148294  
157304: contig of 9011 bp in length  
157305  
157404: gap of unknown length  
157405  
168245: contig of 10841 bp in length.

#### FEATURES

Location/Qualifiers  
1..168245  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
/clone="CH230-129p3"

BASE COUNT 43979 a 35809 c 36510 g 44464 t 7483 others  
ORIGIN

#### Alignment Scores:

Pred. No.: 3.83e-83 Length: 168245  
Score: 1152.00 Matches: 214  
Percent Similarity: 98.22% Conservative: 7  
Best Local Similarity: 95.11% Mismatches: 4  
Query Match: 77.79% Indels: 0  
DB: 2 Gaps: 0

US-09-513-365A-1 (1-280) x AC128348 (1-168245)

Qy 56 TyrglnGluThrLeuLysGluLeuAspValTyrglnLysTyrglnLysGluAsp 75  
Db 100161 TTTTAAAGAACTTAAAGAAATTCATGCTCTATGAAAAATAAGAAAGAGATCAT 100220  
Qy 76 LeuAsnGlnLysLysArgLeuGlnGlnLeuGlnArgAlaLeuIleAsnSerGlnGlu 95  
Db 100221 TCNAACCAAGAAAAACCGCTACGACGATCTCCAGAGAGCATTAATCAATAGCCAGAA 100280

```
QY 96 LeuGlyAspGluLysIleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAla 115
|||||
Db 100281 TTGGGAGATCAAAAGATTGATGTCACACAGATCTGGAATTGCTGAAACCCGAGCC 100340
QY 116 ArgGlnMetLeuHisSerGlnCysPheGlnAspProAlaGluSerGluArgAlaSer 135
|||||
Db 100341 AGACAAATGGAGTGCATTACAGTCTTCCAGGATCCCTGCTGAAAGTGGCGGGCTCA 100400
QY 136 AspLysAlaLysMetAspSerSerGlnProGluArgSerSerArgArgProArgGln 155
|||||
Db 100401 GACAAATCCAAGATGATCCAGTCAACCCGAAAGATCTCCAGAAGACCTCGAAGACAG 100460
QY 156 ArgThrSerGluSerArgAspLeuGluCysHisMetAlaAsnGlyIleGluAspCysAsp 175
|||||
Db 100461 CGGACCAAGTGGAGCGCTGATTTATGTCACATGACAAAGCGGATTTGATGCTCGATGAT 100520
QY 176 GlnProProLysGluLysSerLysSerAlaLysLysLysLysLysLysLysLys 195
|||||
Db 100521 CAGCCACCCGAAAGAAAGAAATCCAAATCCGCCAAGAAAGAAAGCGCTCCAAAGGCCAAG 100580
QY 196 GlnGluArgGluAlaSerProValGluPheAlaIleAspProAsnGluProThrTyrCys 215
|||||
Db 100581 CAGGAGGGAAGCGTCACCTGTTGAGTTGGCATAGATCCCAACGAGGCCACTACTGCT 100640
QY 216 LeuCysAsnGlnValSerTyrGlyGluMetIleGlyCysAspAsnGluCysProile 235
|||||
Db 100641 TTATGTAAACCAAGTGTCTTACGGGAGATGATAGGCTGTGACAAATGAACAGTGTCCATT 100700
QY 236 GluTrpPheHisPheSerCysValSerLeuThrTyrLysProLysGlyLysTrpTyrCys 255
|||||
Db 100701 GAGTGTGTTCACCTTTTCCTGCGTTTCTACCTACCTATAAACCAGGAAAGTGGATGCT 100760
QY 256 ProLysCysArgGlyAspAsnGluLysThrMetAspLysSerThrGluLysThrLysLys 275
|||||
Db 100761 CCAAAGTCAGGGGACAGACAGTGAAGAAACCATGACGACAAAGTACTGAAGACACAAAG 100820
QY 276 AspArgArgSerArg 280
|||||
Db 100821 GAGAGAGAGCGGAGG 100835

RESULT 10
AY014017
LOCUS AY014017 1094 bp mRNA linear VRT 16-DEC-2001
DEFINITION Xenopus laevis p32INGL mRNA, complete cds.
ACCESSION AY014017
VERSION AY014017.1 GI:17864714
KEYWORDS
SOURCE
ORGANISM Xenopus laevis.
Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus.
REFERENCE
1 (bases 1 to 1094)
Wagner,M.J., Gogela-Spehar,M., Skirrow,R.C., Johnston,R.N.,
Riabowol,K. and Helbing,C.C.
Expression of Novel ING Variants Is Regulated by Thyroid Hormone in
the Xenopus laevis Tadpole
J. Biol. Chem. 276 (50), 47013-47020 (2001)
11600495
2 (bases 1 to 1094)
Helbing,C.C., Wagner,M.J., Gogela-Spehar,M., Skirrow,R.C.,
Johnston,R.N. and Riabowol,K.
Direct Submission
Submitted (16-NOV-2000) Biochemistry and Microbiology, University
of Victoria, P.O. Box 3055 Stn. CSC, Victoria, British Columbia V8W
3P6, Canada
FEATURES
source
location/Qualifiers
1..1094
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/tissue_type="brain"
/dev_stage="tadpole"
109..945
CDS
```



	Qy	232 GlnCysProIleGluTrpPheHisSerCysValserLeuthrTyr	247   
	Dd	14651 CAGTGCCCAATGAATGGTTTCACCTTTTCATGTGTTTCACCTTACCTAT	14604
RESULT 12			
AF149820                  AFL149820        840 bp      mRNA     linear      ROD 21-NOV-1999			
LOCUS Mus musculus p33INGl protein mRNA, complete cds.			
DEFINITION AF149820			
ACCESSION AF149820			
VERSION AF149820.1 GI:6456561			
KEYWORDS .			
SOURCE Mus musculus.			
ORGANISM Mus musculus			
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
REFERENCE	1	(bases 1 to 840)	
AUTHORS	Rancourt,D. and Garkavtsev,I.		
TITLE	Structural organization and expression pattern of the murine INGI gene		
JOURNAL	Unpublished		
REFERENCE	2	(bases 1 to 840)	
AUTHORS	Garkavtsev, I V		

```

direct submission
file
JOURNAL Submitted (11-MAY-1999) Genome Therapeutics Corp., 100 Beaver St.,
Waltham, MA 02453, USA
FEATURES Location/Qualifiers
source 1..840
/organism="Mus musculus"
/db_xref="taxon:10090"
CDS 1..840

```

```
/condition="p33ING1 protein"
```

<pre> /db_xref="GI:6456562" /protein_id="AAP09183.1" /translation="MLSPANGQIHLVNYVEDYLDISIEPLFLQRNVSLMREIDAKY OETLLEDYYEKFRNTOGQRRVLHCIOALRISIQSGDEKIQIVSOMVELVENR SQVDVSHEFLEAHODISDGTSGSKGAKDSKSAITOADPKNRKSORRNNEIR NASNHDDHDIJTSCHPKEKATSKKKRSKASPEADPDIPDETPTVCILCNOV SYGEMTCNDCECPTEWFHFSCVGLNHHKPKGWYCPKCRGESEKTMDKALEKSKVERA YNR" </pre>					
BASE COUNT	250 a	216 c	251 g	123 t	
ORIGIN					
Alignment Scores:					
Pred. No.:	9,81e-64	Length:	840		
Score:	880.50	Matches:	166		
Percent Similarity:	77.78%	Conservative:	44		
Best Local Similarity:	61.48%	Mismatches:	49		
Query Match:	59.45%	Indels:	11		
DB:	10	Gaps:	3		
US-09-513-365A-1 (1-280) x AF149820 (1-840)					
Qy	18	GlyGluArgSerArgLeuLeuThrCysTyrValGlnAspTyrLeuCysValGluSer	37		
Db	19	GGGACCATCCACTCGTGTAAC---TATGTGGAGATTACCTGGACTCAATCGAGTCA	75		
Qy	38	LeuProHisAspMetGlnArgAsnValSerValLeuArgGluLeuAspAsnLysTyrGln	57		
Db	76	CTGCCTTTGCACCTGCAGAGGAACGTCCTCGCTGATCGGGGAGATCGACGCCAAATACCAA	135		
Qy	58	GluThrLeuLysGluIleAspValTyrGluLuTyrLysLysGluAspAspleuAsn	77		
Db	136	GAGATCCTCAAGGAGCTGGACCACTACTATGAGAAGTTCCAACGGGAGACACAGCGCAC	195		
Qy	78	GlnLysLysArgLeuGlnGlnLeuLeuGlnArgAlaLeuIleAsnSerGlnGluLeuGly	97		
Db	196	CAGAAGCGGGTACTGCTGCATCCAGAGGCCCTGATCCGCAGCCAGGAGTAGGC	255		
Qy	98	AspGluLysIleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAlaArgGln	117		
Db	256	GATGAGAAGATCCAGATCGTGAGTCAGATGGTGGAGCTGCTGGAGAACCGCAGCAGACAG	315		

Alignment Scores:	
Pred. No.:	9, 81e-64
Score:	880.50
Percent Similarity:	77.78%
Best Local Similarity:	61.48%
Query Match:	95.45%
DB:	10
	5
	3
Length:	840
Matches:	166
Conservative:	44
Mismatches:	49
Indels:	11
Gaps:	3

US-09-513-365A-1 (1-280) X AF149820 (1-840) .

Qy	18	GlyGluuSerArgLeuLeuthrCysTyrValGlnAspTyrIeuGluCysValGluSer	37
Dd	19	GGGAGACCATCCACTGGTGAACT--TATGTGGAGATTACTCGCACTCAATCGAGTGA	75
Qy	38	LeuProHisAspMetGlnArgAsnValSerValLeuArgGluLeuAspAsnLysTyrGln	57
Dd	76	CTGCCTTTTCGACTGCAGAGGAACGCTCTCGTGATCGGGAGATCGACGCCAATAACCAA	135
Qy	58	GluThrLeuLysGluLeuLeuLeuLysTyrGluLysTyrLysGluGluAspAspLeuAsn	77
Dd	136	GAGATCCTCAAGGAGCTGGACGCACTACTATGAGAATTCCAACGGGAGACACACGCCACC	195
Qy	78	GlnLysLysArgGlnGlnLeuLeuGlnArgAlaLeuLeuAsnSercGlnGluLeuGly	97
Dd	196	CAGAMGCCGGGTACTGCATGCATCCAGAGGCCCTGTTCGCGACCCGAGAGCTAGGC	255
Qy	98	AspGluLysIleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAlaArgGln	117
Dd	256	GATGAGAAGATCCAGATCGTAGTCAGATGGTGGAGCTGGTGGAGAACCGCAGACAGACAG	315



```

Qy 118 MetGluLeuHisSerGlnCysPhe-----GlnAspProAlaGluSerGluArgAla 134
Db 316 GTGGACAGTCACGTGGAGCTCTCGAAGACACACAGACATCAGTCAGCGCAGTGGTGGC 375
Qy 135 SerAspLysAlaLysMetAspSerSerGlnProGluArg----- 147
Db 376 AGCGGACAGCGGGCCAGGACAGTCGAAGAGTGAGGCCATCAGAGGAGAGTAAGCGG 435
Qy 148 SerSerArgArgProArgGlnArgThrSerGluSerArgAspLeuCysHisMetAla 167
Db 436 AATAACAAGGCTTCAGAGGAGCAGCAACAATGAGAATCGAGAGACGGCTCGAATAAT 495
Qy 168 AsnGlyLeuGluAspCysAspGlnProProLysGluLysSerLysSerAlaLys 187
Db 496 CAGCAGCATGATGATCATCCTCAGAACGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAG 555
Qy 188 LysLysLysArgSerLysAlaLysGlnGluArgGluAlaSerProValGluPheAlaLe 207
Db 556 AAGAAGAAGCGTCCAGAGGCAAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 615
Qy 208 AspProAsnGluProThrTyrCysLeuCysAsnGlnValSerTyrGlyGluMetIleGly 227
Db 616 GACCCCAAGCAGGCGCAGTACTGTCTGTGCAACAGAGTCTCTACGGGGAGATGATCGG 675
Qy 228 CysAspAsnGluGlnCysProIleGluTrpPheHisPheSerCysValSerLeuThrTyr 247
Db 676 TGTGACAAACAGCAGTACCCATCGAGTGGTTCACCTCTCCTCGGTGGGCTCAACCAT 735
Qy 248 LysProLysGlyLysTyrTyrCysProLysCysArgGlyAspAsnGluLysThrMetAsp 267
Db 736 AAACCAAGGCGAAGTGTACTGCCCAAGTGGCGTGGGAGAGAGAGAGAGAGAGAGAGAG 795
Qy 268 LysSerThrGluLysThrLysLysAspArg 277
Db 796 AAGCCCTTGAGAGTCCAAAGAGAGAGG 825

RESULT 13
BC016573
LOCUS
DEFINITION Mus musculus, inhibitor of growth family, member 1, clone MGC:27881
IMAGE:3496013, mRNA, complete cds.
ACCESSION BC016573
VERSION BC016573.1 GI:16741521
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Strausberg, R.
Direct Submission
Submitted (31-OCT-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amc@bcm.tmc.edu
Gonaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 35 Row: 1 Column: 6.

```

```

FEATURES
source
Location/Qualifiers
1..1584
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="C57BL/6J"
/clone="MGC:27881 IMAGE:3496013"
/tissue_type="Mammary tumor. WAP-TGF alpha model. 7 months
old, gross tissue."
/clone_lib="NCI_CGAP_Mam5"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
389..1228
/codon_start=1
/product="inhibitor of growth family, member 1"
/protein_id="AAH16573.1"
/db_xref="GI:16741522"
/db_xref="LocusID:26356"
/translation="MSPANGEOIHLVNYVEDYLDSTESLPFDLQNVSLMREIDAKY
QETLKELDDYEFKRETDGTQRRVLCQRLIRSEQLGDEKIQIVSQWVELVNR
SRQVDSHELVEAFHQDISDGTGGSGKAGDKSKEAITQADKPNKKRSRRNNRE
NAGNNHDDDDITSGTPKPKAKTSKKKSKAKAEASPADLPIDPNEPTCYLCNQV
SYGEMICNDNDECPIEWFHFSCVGLNHKPKWKYCPKRGSEKTMKALEKSKKERA
YNR"
BASE COUNT 384 a 422 c 517 g 261 t
ORIGIN
Alignment Scores:
Pred. No.: 6,79e-63 Length: 1584
Score: 874.50 Matches: 165
Percent Similarity: 77.41% Conservative: 44
Best Local Similarity: 51.11% Mismatches: 50
Query Match: 59.05% Indels: 11
DB: 10 Gaps: 3
US-09-513-365A-1 (1-280) x BC016573 (1-1584)
Qy 18 GlyGluArgSerArgLeuLeuThrCysTyrValGlnAspTyrLeuGluCysValGluSer 37
Db 407 GGGGAGCAGATCCACCTGGTGAAC---TATGTGGAGGATTACCTGGACTCAATCGAGTCA 463
Qy 38 LeuProHisAspMetGlnArgAsnValSerValLeuArgGluLeuAspAsnLysTyrGln 57
Db 464 CTGCTTTCCACCTGCGAGAGGAGCTCTCGCTGATGGGGAGATCGAGCCCAATACCAA 523
Qy 58 GluThrLysGluLysGluLeuAspValTyrGluLysTyrLysGluAspLeuAsn 77
Db 524 GAGATCTCTGAAGAGCTGGACGACTACTATGAGAAGTTCAAACGGGAGACAGACGCGACC 583
Qy 78 GlnLysLysArgLeuGlnGlnLeuLeuGlnArgAlaLeuLeuAsnSerGlnGluLeuGly 97
Db 584 CAGAAGCGCGGGTACTGCATCTGCATCCAGAGGCGCTGATCCGACGAGGAGCTAGGC 643
Qy 98 AspGluLysIleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAlaArgGln 117
Db 644 GATGAGAAGATCCAGATCGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 703
Qy 118 MetGluLeuHisSerGlnCysPhe-----GlnAspProAlaGluSerGluArgAla 134
Db 704 GTGGACAGTCACGTGGAGCTCTTCGAAGACACACAGGACATCAGTCAGCGCAGTGGTGGC 763
Qy 135 SerAspLysAlaLysMetAspSerSerGlnProGluArg----- 147
Db 764 AGCGGACAGCGGGCCAGGACAGTCAAGAGAGTGGAGGAGTCAAGAGAGAGAGAGAGAG 823
Qy 148 SerSerArgArgProArgGlnArgThrSerGluSerArgAspLeuCysHisMetAla 167
Db 824 AATAACAAGGCTTCAGAGGAGCAGCAACAATGAGAATCGAGAGACGGCTCGAATAAT 883
Qy 168 AsnGlyIleGluAspCysAspGlnProProLysGluLysSerLysSerAlaLys 187
Db 884 CAGCAGCATGATGATCATCCTCAGAGAGCGCCCAAGAGAGAGAGAGAGAGAGAGAGAG 943
Qy 188 LysLysLysArgSerLysAlaLysGlnGluArgGluAlaSerProValGluPheAlaLe 207

```



## AUTHORS

Nagashima,M., Shiseki,M., Miura,K., Hagiwara,K., Linke,S.P.,  
 Pedoux,R., Wang,X.W., Yokota,J., Riabowol,K. and Harris,C.C.  
 DNA damage-inducible gene p33ING2 negatively regulates cell  
 proliferation through acetylation of p53  
 Proc. Natl. Acad. Sci. U.S.A. 98 (17), 9671-9676 (2001)

## MEDLINE

21396501

## PUBMED

11481424

## REFERENCE

2 (bases 1 to 840)

## AUTHORS

## TITLE

Submitted (16-JUL-1998) Laboratory of Human Carcinogenesis,  
 National Cancer Institute, National Institutes of Health, 37  
 Convent Dr., Bldg. 37, Rm. 2C01, Bethesda, MD 20892, USA

## JOURNAL

Location/Qualifiers

## FEATURES

source

## gene

1..840 /organism="Homo sapiens"

## CDS

1..840 /db\_xref="taxon:9606"

1..840 /gene="ING1"

1..840 /codon\_start=1

1..840 /product="p33ING1"

1..840 /protein\_id="AAG12174.1"

1..840 /db\_xref="GI:10039545"

translation="MLSPANGEOQLHVNYYEDYIDSTESLPFDLQNVSLMREIDAKY  
 OEILKELDECYERFSRETGAQRMLHCYORALIRSOELGDEKIOIVSOMVELVNR  
 TROVDSHVELFEAGDELDTAGNSKAGADRPKGEAAQADKPSKRSRQRNNEFE  
 NASNNHDDGAGSTPEKKAKTSKKKSKAKAEREASPADLPIDNPEYCLCNQV  
 SYGEMIGCDNDEGPTIEWFHFCVGLNHKPKGWYCPKRCGENEKTMDKALEKSKKERA  
 YNR"

BASE COUNT 218 a 239 c 277 g 106 t

## ORIGIN

## Alignment Scores:

Pred. No.: 1.62e-61 Length: 840  
 Score: 853.50 Matches: 160  
 Percent Similarity: 76.30% Conservative: 46  
 Best Local Similarity: 59.26% Mismatches: 53  
 Query Match: 57.63% Indels: 11  
 DB: 9 Gaps: 3

US-09-513-365A-1 (1-280) x AF078835 (1-840)

Qy 18 GlyGluArgSerArgLeuThrCysTyrValGlnAspTyrLeuGluCysValGluSer 37  
 Db 19 GGGGAGGAGCTCCACCTGGTGAC--TATGTGAGGAGCTACCTGGAGCTCATCGAGTCC 75  
 Qy 38 LeuProHisAspMetGlnArgAsnValSerValLeuArgGluLeuAspAsnLysTyrGln 57  
 Db 76 CTGCCTTTCGACTTGCAGAGAAATGTCTCGTGTATGGGAGATCGACGCGAAATACCAA 135  
 Qy 58 GluThrLeuLysGluLeuAspValTyrGluLysTyrLysLysGluAspLeuAsn 77  
 Db 136 GAGATCCTGAAGGAGCTAGACGAGTGTCTACGCGCTTCAGTCGCGAGACAGCGGGCG 195  
 Qy 78 GlnLysLysArgLeuGlnLeuLeuGlnArgAlaLeuLeuAsnSerGlnLeuGly 97  
 Db 196 CAGAGCGGGGATGTGCTGTGTCAGCGCGCTGATCCGCGAGAGCTGGG 255  
 Qy .98 AspGluLysIleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAlaArgGln 117  
 Db 256 GACGAGAAGATCCAGATCGTGGCAGACAGTGTGGAGCTGTGGAGAACCCGACGCGGCGAG 315  
 Qy 118 MetGluLeuHisSerGlnCysPheGlnAspProAlaGlu---SerGluArgAlaSerAsp 136  
 Db 316 GTGGACAGCCACCTGGAGCTGTTCGAGCGCGACGAGGAGTGGCGCGACAGCGGGCGAAC 375  
 Qy 137 LysAlaLysMetAspSerSerGlnPro-----GluArg 147  
 Db 376 AGCGGCAAGCTGGCGGAGACGCCCAAGAGCGGCGGCGGAGCGGAGCTGACAGGCC 435  
 Qy 148 SerSerArgArgProArgGlnArgThrSerGluSerArgAspLeuCysHisMetAla 167

Search completed: June 20, 2003, 03:27:45  
 Job time : 2265 secs

Db 436 AACAGCAAGCGCTCACGGCGCGCAGCGCAACAACAGAGAACCGTGAGAACGGCTCCAGCAAC 495  
 Qy 168 AsnGlyIleGluAspCysAspGlnProProLysGluLysLysSerLysSerLysSer 187  
 Db 496 CACGACCACGACGCGCGCTCGGGCACACCCCAAGAGAGAGAGGCGCAAGACCTCCAAG 555  
 Qy 188 LysLysLysArgSerLysAlaLysGlnGluArgGluAlaSerProValGluPheAlaIle 207  
 Db 556 AAGAAGAAGCGCTCCAAGGCCAAGGCGGAGGAGGCGTCCCTGCGGACCTCCCATC 615  
 Qy 208 AspProAsnGluProThrTyrCysLeuCysAsnGlnValSerTyrGlyGluMetIleGly 227  
 Db 616 GACCCCAACGAACCCACGTACTGTCTGTGCAACAGGTCTCTATGGGAGATGATCGGC 675  
 Qy 228 CysAspAsnGluGlnCysProIleGluTrpPheHisPheSerCysValSerLeuThrTyr 247  
 Db 676 TGGACAAACGACGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 735  
 Qy 248 LysProLysGlyLysTyrTyrCysProLysCysArgGlyAspAsnGluLysThrMetAsp 267  
 Db 736 AAACCCAAAGGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 795  
 Qy 268 LysSerThrGluLysThrLysLysAspArg 277  
 Db 796 AAAGCCCTGGAGAAATCCAAAAAGAGAGG 825



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 20, 2003, 02:42:14 ; Search time 68 seconds  
(without alignments)  
1262.787 Million cell updates/sec

Title: US-09-513-365A-1

Perfect score: 1481

Sequence: 1 MLGQQQQQLYSSAALLTGER.....DNEKTMDSKTEKTKDRRSR 280

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 16.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model ; -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool/US09513365/runat\_16062003\_124743\_16468/app\_query.fasta\_1.455  
-DB=Issued\_Patents\_NA -Qfmt=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09513365.qcgn\_1.1\_61 @runat\_16062003\_124743\_16468 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents.NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1481	100.0	840	US-09-601-478-6	Sequence 6, Appli
2	1481	100.0	1078	US-09-601-478-7	Sequence 7, Appli
3	1334	90.1	1154	US-09-484-970B-81	Sequence 81, Appli
4	853.5	57.6	873	US-09-006-783A-4	Sequence 4, Appli
5	728	49.2	1902	US-09-258-257-1	Sequence 1, Appli
6	728	49.2	1902	US-09-258-371-1	Sequence 1, Appli
7	728	49.2	1902	US-08-569-721A-1	Sequence 1, Appli
8	728	49.2	1902	US-08-751-230-1	Sequence 1, Appli
9	728	49.2	1902	US-09-499-082-1	Sequence 1, Appli
10	728	49.2	1902	US-09-258-372-1	Sequence 1, Appli
11	728	49.2	2061	US-09-258-371-9	Sequence 9, Appli
12	728	49.2	2061	US-08-751-230-9	Sequence 9, Appli

13	728	49.2	2061	3	US-09-499-082-9	Sequence 9, Appli
14	728	49.2	2061	4	US-09-258-372-9	Sequence 9, Appli
15	728	49.2	2061	4	US-09-159-871-1	Sequence 1, Appli
16	722	48.8	2061	4	US-09-006-783A-2	Sequence 2, Appli
17	662	44.7	633	4	US-09-006-783A-6	Sequence 6, Appli
18	400	27.0	699	3	US-09-195-286-2	Sequence 2, Appli
19	396.5	26.8	678	3	US-09-195-286-3	Sequence 3, Appli
20	314	21.2	451	4	US-09-370-838-146	Sequence 146, App
21	172.5	11.6	8931	3	US-09-051-019-1	Sequence 1, Appli
22	127.5	8.6	1820	4	US-08-845-258-17	Sequence 17, Appli
23	127.5	8.6	1820	4	US-08-845-258-37	Sequence 37, Appli
24	127.5	8.6	1820	4	US-08-990-571-17	Sequence 17, Appli
25	127.5	8.6	1820	4	US-08-990-571-37	Sequence 37, Appli
26	127.5	8.6	1820	4	US-08-723-142A-17	Sequence 17, Appli
27	127.5	8.6	1820	4	US-08-723-142A-37	Sequence 37, Appli
28	127.5	8.6	1820	4	US-09-528-784A-17	Sequence 17, Appli
29	127.5	8.6	1820	4	US-09-528-784A-37	Sequence 37, Appli
30	127.5	8.6	2001	4	US-09-528-784A-84	Sequence 84, Appli
31	127.5	8.6	3402	4	US-09-528-784A-86	Sequence 86, Appli
32	125	8.4	6328	4	US-08-913-832A-1	Sequence 1, Appli
33	125	8.4	6328	4	US-09-249-181A-1	Sequence 1, Appli
34	122	8.2	8600	4	US-09-457-708-1	Sequence 13, Appli
35	120	8.1	1772	2	US-08-960-022-13	Patent No. 5273901
36	116.5	7.9	543	6	5273901-6	Sequence 93, Appli
37	116.5	7.9	9551	1	US-08-056-200-93	Sequence 93, Appli
38	116.5	7.9	9551	2	US-08-800-644-93	Sequence 401, App
39	115	7.8	812	4	US-08-998-416-401	Sequence 2, Appli
40	114	7.7	1053	5	PCT-US91-06418-2	Sequence 96, Appli
41	114	7.7	1621	2	US-08-933-750C-96	Sequence 96, Appli
42	114	7.7	1621	3	US-09-234-613-96	Sequence 195, App
43	113.5	7.7	6693	4	US-08-961-527-195	Sequence 1, Appli
44	113	7.6	4181	4	US-09-393-569-1	Sequence 38, Appli
45	112.5	7.6	1864	1	US-08-454-720A-38	

#### ALIGNMENTS

##### RESULT 1

US-09-601-478-6

; Sequence 6, Application US/09601478  
; Patent No. 6403785  
; GENERAL INFORMATION:  
; TITLE OF INVENTION: Human TSC403 gene and human ING1L gene  
; APPLICANT: Ostuka Pharmaceutical Co., Ltd.  
; FILE REFERENCE: 060193  
; CURRENT APPLICATION NUMBER: US/09/601.478  
; CURRENT FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: JP H10-134679  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: JP H10-73234  
; PRIOR FILING DATE: 1998-03-05  
; PRIOR APPLICATION NUMBER: JP H10-38133  
; PRIOR FILING DATE: 1998-02-03  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 840  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: human embryonic brain cDNA library  
US-09-601-478-6

Alignment Scores:  
Pred. No.: 3.31e-156 Length: 840  
Score: 1481.00 Matches: 280  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-513-365A-1 (1-280) x US-09-601-478-6 (1-840)

```
QY 1 MetLeuGlyGlnGlnGlnGlnLeuTyrSerSerAlaAlaLeuLeuThrGlyGluArg 20
Db 1 ATGTTAGGCAGCAGCAGCAGCACTGTACTCGTCGCCCGCGCTCTGACCGGGAGCGG 60
QY 21 SerArgLeuLeuThrCysTyrValGlnAspTyrLeuGluCysValGlnSerLeuProHis 40
Db 61 AGCCGGCTCTCACCCTGCTACGTGCGAGGACTACCTTGAGTGGCTGGAGTCGCTGCCCCAC 120
QY 41 AspMetGlnArgAsnValSerValLeuArgGluLeuAspAsnLysTyrGlnGluThrLeu 60
Db 121 GACATCCAGAGAACGCTCTGCTGCGAGAGCTGACCAACAAATATCAAGAACGTTA 180
QY 61 LysGluIleAspAspValTyrGluLysTyrLysGluAspAspLeuAsnGlnLysLys 80
Db 181 AAGGAATATGATGATCTACGAAATATAGAAAGAGATGATTAAACAGAGAA 240
QY 81 ArgLeuGlnGlnLeuLeuGlnArgAlaLeuIleAsnSerGlnGluLeuGlyAspGluLys 100
Db 241 CGTCTACAGCAGCTCTCCAGAGACCTTAATTAAGTCAAGAAATGGGAGATGA 300
QY 101 IleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAlaArgGlnMetGluLeu 120
Db 301 ATACAGATGTTTACACAAATGCTCGAATGTTGGAAATCGGSCAAGACAAATGGAGTTA 360
QY 121 HisSerGlnCysPheGlnAspProAlaGluSerGluArgAlaSerAspLysAlaLysMet 140
Db 361 CACTCACAGTGTTCACAGATCCTGCTGAAAGTGACAGAGCCTCAGATAAAGCAAGATG 420
QY 141 AspSerGlnProGluArgSerArgArgProArgArgGlnArgThrSerGluSer 160
Db 421 GATTCAGCCACACAGAAAGATCTTCAAGAGACCCCGCAGCGGAGCCAGTGAAGC 480
QY 161 ArgAspLeuCysHisMetAlaAsnGlyIleGluAspCysAspAspGlnProProLysGlu 180
Db 481 CGTGATTTATGTCACATGCGAAATGGATGTAAGACTGTGATGATCAGCCACCTAAAGAA 540
QY 181 LysLysSerLysSerAlaLysLysLysLysArgSerLysAlaLysGlnGluArgGluAla 200
Db 541 AAGAAATCAAGTCACGAAAGAAAGAAAGAACGCTCCAGGCCAAGCAGAAAGGGAAGCT 600
QY 201 SerProValGluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnVal 220
Db 601 TCACCTGTTGAGTTGTCATAGATCCTAATGAACTACATCTGCTTATGCAACCAAGTG 660
QY 221 SerTyrGlyGluMetIleGlyCysAspAsnGluGlnCysProIleGluTrpPheHisPhe 240
Db 661 TCTTATGGGAGATGATGATGATGACATGAACAGTGTCCAATTCATGTTTCACTTT 720
QY 241 SerCysValSerLeuThrTyrLysProLysGlyLysTyrTyrCysProLysCysArgGly 260
Db 721 TCATGTTTTCATCTTACCTATAACCAAGGGAATGTTATGTCCTCAAGTGCAGGGA 780
QY 261 AspAsnGluLysThrMetAspLysSerThrGluLysThrLysLysAspArgArgSerArg 280
Db 781 GATAATGAGAAACAAATGACAAAGTACTGAAAGACAAAGAAAGATAGAGATGCGAGG 840
```

## RESULT 2

```
US-09-601-478-7
; Sequence 7, Application US/09601478
; Patent No. 6403785
; GENERAL INFORMATION:
; APPLICANT: Ostuka Pharmaceutical Co., Ltd.
; TITLE OF INVENTION: Human TSC403 gene and human ING1L gene
; FILE REFERENCE: Q60193
; CURRENT APPLICATION NUMBER: US/09/601,478
; CURRENT FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: JP H10-134679
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: JP H10-73234
; PRIOR FILING DATE: 1998-03-05
; PRIOR APPLICATION NUMBER: JP H10-38133
; PRIOR FILING DATE: 1998-02-03
; NUMBER OF SEQ ID NOS: 11
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 1078
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human embryonic brain cDNA library
; NAME/KEY: CDS
; LOCATION: (92)..(931)
US-09-601-478-7
```

## Alignment Scores:

```
Pred. No.: 4,78e-156 Length: 1078
Score: 1481.00 Matches: 280
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0
```

US-09-513-365A-1 (1-280) x US-09-601-478-7 (1-1078)

```
QY 1 MetLeuGlyGlnGlnGlnGlnLeuTyrSerSerAlaAlaLeuLeuThrGlyGluArg 20
Db 92 ATGTTAGGCAGCAGCAGCAGCACTGTACTCGTCGCCCGCGCTCTGACCGGGAGCGG 151
QY 21 SerArgLeuLeuThrCysTyrValGlnAspTyrLeuGluCysValGlnSerLeuProHis 40
Db 152 AGCCGGCTCTCACCCTGCTACGTGCGAGGACTACCTTGAGTGGCTGGAGTCGCTGCCCCAC 211
QY 41 AspMetGlnArgAsnValSerValLeuArgGluLeuAspAsnLysTyrGlnGluThrLeu 60
Db 212 GACATCCAGAGAACGCTGCTGCTGCGAGAGCTGGACAAACAAATATCAAGAACGTTA 271
QY 61 LysGluIleAspAspValTyrGluLysTyrLysGluAspAspLeuAsnGlnLysLys 80
Db 272 AAGGAATATGATGATGCTACGAAATATAGAAAGAGATGATTAAACAGAGAA 331
QY 81 ArgLeuGlnGlnLeuLeuGlnArgAlaLeuIleAsnSerGlnGluLeuGlyAspGluLys 100
Db 332 CGTCTACAGCAGCTCTCCAGAGACCTTAATTAAGTCAAGAAATGGGAGATGA 391
QY 101 IleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAlaArgGlnMetGluLeu 120
Db 392 ATACAGATGTTTACACAAATGCTGGAATTTGGGAAATCGGCAAGACAAATGGAGTTA 451
QY 121 HisSerGlnCysPheGlnAspProAlaGluSerGluArgAlaSerAspLysAlaLysMet 140
Db 452 CACTCACAGTGTTCACAGATCCTGCTGAAAGTGAACGAGCCTCAGATTAAGCAAGATG 511
QY 141 AspSerGlnProGluArgSerArgArgProArgArgGlnArgThrSerGluSer 160
Db 512 GATTCAGCAACCAAGAAAGATCTTCAAGAGACCCCGCAGCGGAGCCAGTGAAGC 571
QY 161 ArgAspLeuCysHisMetAlaAsnGlyIleGluAspCysAspGlnProProLysGlu 180
Db 572 CCGTATTTATGTCATGCGAAATGGGATTTGAAGACTGTGATGATCAGCCACTAAAGAA 631
QY 181 LysLysSerLysSerAlaLysLysLysLysArgSerLysAlaLysGlnGluArgGluAla 200
Db 632 AAGAAATCCAGTCAGCAAGAAAGAAAGAAACCTCCAGGCCAAGCAGAGAGGAGGAGCT 691
QY 201 SerProValGluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnVal 220
Db 692 TCACCTGTTGAGTTTGCATAGATCCTAATGAACCTACATCTGCTTATGCAACCAAGTG 751
QY 221 SerTyrGlyGluMetIleGlyCysAspAsnGluGlnCysProIleGluTrpPheHisPhe 240
Db 752 TCTTATGGGAGATGATGAGGATGTGACAATGAACAGTGTCCAATTAATGTTTCACTTT 811
QY 241 SerCysValSerLeuThrTyrLysProLysGlyLysTyrTyrCysProLysCysArgGly 260
Db 812 TCATGTTTTCATCTTACCTATAACCAAGGGAATGTTATGTCCTCAAGTGCAGGGA 871
```



STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 7..813  
US-09-006-783A-4

## Alignment Scores:

Pred. No.: 3,74e-86 Length: 873  
Score: 853.50 Matches: 161  
Percent Similarity: 75.74% Conservative: 45  
Best Local Similarity: 59.19% Mismatches: 51  
Query Match: 57.63% Indels: 15  
DB: Gaps: 3

US-09-513-365A-1 (1-280) x US-09-006-783A-4 (1-873)

QY 18 GlyLysArgSerArgLeuLeuThrCysTyrValGlnAspTyrLeuGluCysValGluSer 37  
DB 25 GGGGAGCAGCTCCACCTGGTGAAC---TATGTGGAGGACTACCTGGACTCCATCGAGTCC 81  
QY 38 LeuProHisAspMetGlnArgAnValSerValLeuArgGluLeuAspAsnLysTyrGln 57  
DB 82 CTGCTTTCACCTGACATGACAGAAATGCTCCTGATCGGGAGATCGACCGGAATACCAA 141  
QY 58 GluThrLeuLysGluLeuAspValTyrGluLysTyrLysGluAspLeuAsn 77  
DB 142 GAGATCTGAAGGAGCTAGACGAGTCTACGAGCGCTTCAGTCGCCGAGACAGCGGGCG 201  
QY 78 GlnLysLysArgLeuGlnGlnLeuLeuGlnArgAlaLeuLeuAsnSerGlnGluLeuGly 97  
DB 202 CAGAAGCGCGATGCTGCTGTCGACGCGCGCTGATCGCAGCGAGGAGTGGCG 261  
QY 98 AspGluLysIleGlnLeuValThrGlnMetLeuLeuValGluAsnArgAlaArgGln 117  
DB 262 GACGAGATCAGATCGTGTGACGAGTGGTGGAGTGGTGGAGAACCGCAGCGGCGAG 321  
QY 118 MetGluLeuHisSerGlnCysPheGlnAspProAlaGlu----- 130  
DB 322 GTGGACAGCCACGTGGAGTGTTCGAGCGCGCAGCAGAGCTGGCGGCACACAGTGGCGAAC 381  
QY 131 -----SerGluArgAlaSerAspLysAlaLysMetAspSerGlnPro 145  
DB 382 AGCGGCAAGGTGGCGCGGACAGCGCCCAATGCGATGCGGTAGCGAGTCTGACAAGCC 441  
QY 146 GluArgSerArgArgProArgGlnArgThrSerGluSerArgAspLeuCysHis 165  
DB 442 -----AACAGCAAGCGCTCACGGCGGCGAGCGCAACAGAGAACCGTGAGAACGCGTCC 495  
QY 166 MetAlaAsnGlyIleGluAspCysAspGlnProLysGluLysLysSerLysSer 185  
DB 496 AGCAACACGACGACGACGCGCGCTCGGGCACACCCCAAGAGAGAGAACGCAAGACC 555  
QY 186 AlaLysLysLysLysArgSerLysAlaLysGlnGlnArgGluAlaSerProValGluPhe 205  
DB 556 TCCAGAGAAAGAGCGCTCCAGGCAAGCGGCGAGAGCGGCTCCCTGCGCACCTC 615  
QY 206 AlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnValSerTyrGlyGluMet 225  
DB 616 CCCATGACCCCAACGACACGACTACTCTGTGTGCAACAGCTCCTATGGGAGATG 675  
QY 226 IleGlyCysAspAsnGluGlnCysProIleGluTrpPheHisPheSerCysValSerLeu 245  
DB 676 ATCGGTGCGACAAACAGCAGTGCCTCATCGAGTGGTTCCTCTCTGCTGCTGGGCGTC 735  
QY 246 ThrTyrLysProLysGlyLysTrpTyrCysProLysCysArgGlyAspAsnGluLysThr 265  
DB 736 AATCATAAACCAAGGCAAGTGGTACTCTCCAAAGTCCCGGGGAGAACAGAGACG 795  
QY 266 MetAspLysSerThrGluLysThrLysLysAspArg 277  
DB 796 ATGGACAAAGCCCTGGAGAAATCCAAAAAAGAGAGG 831

## RESULT 5

US-09-258-257-1  
Sequence 1, Application US/09258257  
Patent No. 5965398  
GENERAL INFORMATION:  
APPLICANT: GARKAVTSEV, Igor  
APPLICANT: RIABOWOL, Karl  
TITLE OF INVENTION: DNA SEQUENCE ENCODING A TUMOR  
SUPPRESSOR GENE  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burns, Doane, Swecker & Mathis  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/258,257  
FILING DATE:  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/569,721  
FILING DATE: 08-DEC-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Mooi, Leslie A.  
REGISTRATION NUMBER: 37,047  
REFERENCE/DOCKET NUMBER: 028722-128  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 854-7400  
TELEFAX: (650) 854-8275  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1902 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 109..738  
US-09-258-257-1

Alignment Scores:  
Pred. No.: 1.2e-71 Length: 1902  
Score: 728.00 Matches: 136  
Percent Similarity: 74.79% Conservative: 39  
Best Local Similarity: 58.12% Mismatches: 45  
Query Match: 49.16% Indels: 14  
DB: Gaps: 2

US-09-513-365A-1 (1-280) x US-09-258-257-1 (1-1902)

QY 56 TyrGlnGluThrLeuLysGluIleAspValTyrGluLysTyrLysLysGluAspAsp 75  
DB 31 TGGAAACAGATCTCTAAGGAGGTAGACGAGTCTACGAGCGCTTCTAGTCGCGAGACAGAC 90  
QY 76 LeuAsnGlnLysLysArgLeuGlnGlnLeuLeuGlnArgAlaLeuIleAsnSerGlnGlu 95  
DB 91 GGGGGCCAGAACGGCGGATGTGCTGCTGTGACGCGCGCTGATCCGACCCAGAG 150  
QY 96 LeuGlyAspLysIleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAla 115  
DB 151 CTGGGCGACGAGAGATCCAGATCGTGACCGAGATGGTGAGCTGGTGGAGAACCCAGC 210  
QY 116 ArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGlu----- 130  
DB 116 ArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGlu----- 130



Db 211 CGCAGGTGACACGCCACGTGAGCTTTCAGCGCAGCAGGAGCTGGCGACACACAGTG 270  
QY 131 -----SerGluArgAlaSerAspLysAlaLysMetAspSerSer 143  
Db 271 GGCACAGCGGCAAGGTGGCGGACGAGCCCAATGGCGATGGCGTAGCGCATCTGAC 330  
QY 144 GlnProGluArgSerSerArgProArgArgGlnArgThrSerGluSerArgAspLeu 163  
Db 331 AAGCCC-----AACACGAAGCGCTACGCGCGCAGCGCAACACAGAGACCGTGAGAAC 384  
QY 164 CysHisMetAlaAsnGlyIleGluAspCysAspAspGlnProProLysGluLysLysSer 183  
Db 385 GCCTCCAGCAACCAACACGACGAGCGCGCTCGGGCAGACCCCAAGAGAGAGAGGCC 444  
QY 184 LysSerAlaLysLysLysArgSerLysAlaLysGlnGluArgGluAlaSerProVal 203  
Db 445 AAGACCTCCAAAGAGAGAGCGCTCCAGGCCAAGCGCGAGGAGGAGGCTCCCTGCC 504  
QY 204 GluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnValSerTyrGly 223  
Db 505 GACCTCCCATCGACCCCAACGACCAACCTACTGTCTGTGCAACCAAGGTCTCTATGGG 564  
QY 224 GluMetIleGlyCysAspAsnGluGlnCysProIleGluTrpPheHisPheSerCysVal 243  
Db 565 GAGATGATCGGCTCGGCAACGACGAGTGCCTCCATCGAGTGTTCACCTTCTCGTGGGTG 624  
QY 244 SerLeuThrTyrLysProLysGlyLysTyrTyrCysProLysCysArgGlyAspAsnGlu 263  
Db 625 GGGCTCAATCAATAACCAAGGCAAGTGTGTCTCCCAAGTGCCTCCAGTGGCGGGAGACGAG 684  
QY 264 LysThrMetAspLysSerThrGluLysThrLysLysAspArg 277  
Db 685 AAGACCATGACAAAGCGCTGGAGAAATCCAAAAAGAGAGAG 726

## RESULT 6

US-09-258-371-1  
; Sequence 1, Application US/09258371  
; Patent No. 5986078  
; GENERAL INFORMATION:  
; APPLICANT: Garkavtsev, Igor  
; APPLICANT: Riabowol, Karl  
; TITLE OF INVENTION: DNA SEQUENCE ENCODING THE TUMOR  
; TITLE OF INVENTION: SUPPRESSOR GENE INGI  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Burns, Doane, Swecker & Mathis  
; STREET: 699 Prince Street  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-1404  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/258,371  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/751,230  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mool, Leslie A. 37, 047  
; REGISTRATION NUMBER: 028722-144  
; REFERENCE/DOCKET NUMBER:  
; TELEPHONE: 415-854-7400  
; TELEFAX: 415-854-8275  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1902 base pairs

; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 109..741  
US-09-258-371-1

## Alignment Scores:

Pred. No.: 1,2e-71 Length: 1902  
Score: 728.00 Matches: 136  
Percent Similarity: 74.79% Conservative: 39  
Best Local Similarity: 58.12% Mismatches: 45  
Query Match: 49.16% Indels: 14  
DB: 2 Gaps: 2

US-09-513-365A-1 (1-280) x US-09-258-371-1 (1-1902)

QY 56 TyrGlnGluThrLeuLysGluIleAspAspValTyrGluLysTyrLysLysGluAspAsp 75  
Db 31 TGGAAACAGATCTCTGAAGGAGTAGACAGTGTACGAGCGCTTCAGTCGGCGACACAGAC 90  
QY 76 LeuAsnGlnLysLysArgLeuGlnGlnLeuLeuArgAlaLeuIleAsnSerGlnGlu 95  
Db 91 GGGCGCGAGAAGCGCGGATGCTGCACTGTGTGCGAGCGCGCTGTATCCGCGACGAGGAG 150  
QY 96 LeuGlyAspGluLysIleGlnIleValThrGlnMetLeuGluValGluAsnArgAla 115  
Db 151 CTGGCGCACGAGAGATCCAGATCGTGCAGAGTGTGGAGCTGGTGGAGAACCGCGACG 210  
QY 116 ArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGlu----- 130  
Db 211 CGCGAGGTGGACACCCACCTGGAGCTGTTCCAGGCGCAGCAGGAGCTGGCGCACACAGTG 270  
QY 131 -----SerGluArgAlaSerAspLysAlaLysMetAspSerSer 143  
Db 271 GGCACACGCGCAAGTTGGCGGACAGGCCCAATGGCGATGGGTAGCGCATCTCTAC 330  
QY 144 GlnProGluArgSerSerArgProArgArgGlnArgThrSerGluSerArgAspLeu 163  
Db 331 AAGCCC-----AACACGAAGCGCTCACGGCGGCGAGCAACACAGAGACCGTGAGAAC 384  
QY 164 CysHisMetAlaAsnGlyIleGluAspCysAspAspGlnProProLysGluLysLysSer 183  
Db 385 CGGTCCAGCAACCAACGACGACGAGCGCGCTCGGGCAGACCCCAAGAGAGAGAGGCC 444  
QY 184 LysSerAlaLysLysLysArgSerLysAlaLysGlnGluArgGluAlaSerProVal 203  
Db 445 AAGACCTCCAAAGAGAGAGCGCTCCAAAGGCCAAGCGCGAGGAGCGGCTCCCTGCC 504  
QY 204 GluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnValSerTyrGly 223  
Db 505 GACCTCCCATCGACCCCAACGACCAACCTACTGTCTGTGCAACCAAGGTCTCTATGGG 564  
QY 224 GluMetIleGlyCysAspAsnGluGlnCysProIleGluTrpPheHisPheSerCysVal 243  
Db 565 GAGATGATCGGCTCGGCAACGACGAGTGCCTCCATCGAGTGTTCACCTTCTCGTGGGTG 624  
QY 244 SerLeuThrTyrLysProLysGlyLysTyrTyrCysProLysCysArgGlyAspAsnGlu 263  
Db 625 GGGCTCAATCAATAACCAAGGCAAGTGTGTCTCCCAAGTGCCTCCAGTGGCGGGAGAGAG 684  
QY 264 LysThrMetAspLysSerThrGluLysThrLysLysAspArg 277  
Db 685 AAGACCATGACAAAGCGCTGGAGAAATCCAAAAAGAGAGAG 726

## RESULT 7

US-08-569-721a-1  
; Sequence 1, Application US/08569721a  
; Patent No. 6037121  
; GENERAL INFORMATION:  
; APPLICANT: GARKAVTSEV, Igor



Pred. No.: 1.2e-71 Length: 1902  
Score: 728.00 Matches: 136  
Percent Similarity: 74.79% Conservative: 39  
Best Local Similarity: 58.12% Mismatches: 45  
Query Match: 49.16% Indels: 14  
DB: 3 Gaps: 2

US-09-513-365A-1 (1-280) x US-08-751-230-1 (1-1902)

QY 56 TyrGlnGluThrLeuLysGluLeuAspValTyrGluLysTyrLysGluAsp 75  
DB 31 TGAACACATCTGAGAGAGTAGACGAGTCTACGAGCGCTTCAGTCCGAGACAGAC 90  
QY 76 LeuAsnGlnLysLysArgLeuGlnLeuLeuArgAlaLeuLeuAsnSerGlnGlu 95  
DB 91 GGGGGCAGACGCGGGATGTCACACTGTGTGACGCGCGCTGATCCGACGAGAG 150  
QY 96 LeuGlyAspGluLysIleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAla 115  
DB 151 CTGGGGCAGCAGAAAGATCCAGATCGGTGACGACGATGTGTGAGTGTGTGAGAACCCGACG 210  
QY 116 ArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGlu----- 130  
DB 211 CGGAGGTGGACGACCGCTGAGTGTTCGAGGCGCAGCAGAGTGGGGCAGACAGTG 270  
QY 131 -----SerGluArgAlaSerAspLysAlaLysMetAspSerSer 143  
DB 271 GCGAACAGCGCAAGTGTGGCGGACAGCGCCCAATGGCGATCGGTACGCGAGTGTAC 330  
QY 144 GlnProGluArgSerSerArgProArgGlnArgThrSerGluSerArgAspLeu 163  
DB 331 AAGCCC-----AACAGCAAGCGCTCAGCGGCGCAGCGCAACAGAGAACCGTGAGAAC 384  
QY 164 CysHisMetAlaAsnGlyIleGluAspCysAspGlnProLysGluLysSer 183  
DB 385 GGTTCAGCAACACACACGACGCGCGCTGGGCGACACCGCCCAAGGAGAGAGGCC 444  
QY 184 LysSerAlaLysLysLysArgSerLysAlaLysGlnGluArgGluAlaSerProVal 203  
DB 445 AAGACTCCCAAGAAAGAGCGCTCCAGAGCGCAAGCGGAGGAGCGCTCCCTGCC 504  
QY 204 GluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnValSerTyrGly 223  
DB 505 GACCTCCCATCGACCCCAACAGCAACCGCTACTGTGTGCAACGAGTCTCCATGGG 564  
QY 224 GluMetIleGlyCysAspAsnGluGlnCysProIleGluTrpPheHisPheSerCysVal 243  
DB 565 GAGATGATCGGTGGACACGACGAGTCCCGCTGAGTGTTCACCTTCGTGCGGTG 624  
QY 244 SerLeuThrTyrLysProLysGlyLysTrpTyrCysProLysCysArgGlyAspAsnGlu 263  
DB 625 GGGCTCAATCAATAAACCAAGGCAAGTGTACTGTCCCACTGCGGGGGGAGAACGAG 684  
QY 264 LysThrMetAspLysSerThrGluLysThrLysLysAspArg 277  
DB 685 AAGACATGGACAAAGCCCTGGAGAAATCCAAAAAGAGAGG 726

## RESULT 9

US-09-499-082-1  
; Sequence 1, Application US/09499082  
; Patent No. 6143522  
; GENERAL INFORMATION:  
; APPLICANT: Helbing, Caren C.  
; APPLICANT: Riabowol, Karl  
; APPLICANT: Johnston, Randall N.  
; APPLICANT: Garkavtsev, Igor  
; TITLE OF INVENTION: METHODS OF MODULATING APOPTOSIS  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Burns, Doane, Swecker & Mathis  
; STREET: 699 Prince Street  
; CITY: Alexandria  
; STATE: VA

COUNTRY: USA  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/499,082  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/828,158  
FILING DATE: 27-MAR-1997  
APPLICATION NUMBER: US 08/751230  
FILING DATE: 15-NOV-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/569721  
FILING DATE: 08-DEC-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Mooi, Leslie A.  
REGISTRATION NUMBER: 37,047  
REFERENCE/DOCKET NUMBER: 028722-148  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-854-7400  
TELEFAX: 650-854-8275  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1902 base pairs  
TYPE: nucleic acid  
STRAINEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 109..741  
US-09-499-082-1

Alignment Scores:  
Pred. No.: 1.2e-71 Length: 1902  
Score: 728.00 Matches: 136  
Percent Similarity: 74.79% Conservative: 39  
Best Local Similarity: 58.12% Mismatches: 45  
Query Match: 49.16% Indels: 14  
DB: 3 Gaps: 2

US-09-513-365A-1 (1-280) x US-09-499-082-1 (1-1902)

QY 56 TyrGlnGluThrLeuLysGluLeuAspValTyrGluLysTyrLysGluAsp 75  
DB 31 TGAACACATCTGAGAGAGTAGACGAGTCTACGAGCGCTTCAGTCCGAGACAGAC 90  
QY 76 LeuAsnGlnLysLysArgLeuGlnLeuLeuArgAlaLeuLeuAsnSerGlnGlu 95  
DB 91 GGGGGCAGACGCGGGATGTCACACTGTGTGACGCGCGCTGATCCGACGAGAG 150  
QY 96 LeuGlyAspGluLysIleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAla 115  
DB 151 CTGGGGCAGCAGAAAGATCCAGATCGGTGACGACGATGTGTGAGTGTGTGAGAACCCGACG 210  
QY 116 ArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGlu----- 130  
DB 211 CGGAGGTGGACGACCGCTGAGTGTTCGAGGCGCAGCAGAGTGGGGCAGACAGTG 270  
QY 131 -----SerGluArgAlaSerAspLysAlaLysMetAspSerSer 143  
DB 271 GCGAACAGCGCAAGTGTGGCGGACAGCGCCCAATGGCGATCGGTACGCGAGTGTAC 330  
QY 144 GlnProGluArgSerSerArgProArgGlnArgThrSerGluSerArgAspLeu 163  
DB 331 AAGCCC-----AACAGCAAGCGCTCAGCGGCGCAGCGCAACAGAGAACCGTGAGAAC 384  
QY 164 CysHisMetAlaAsnGlyIleGluAspCysAspGlnProLysGluLysSer 183



COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/258,371  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/751,230  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Mooi, Leslie A.  
REGISTRATION NUMBER: 37,047  
REFERENCE/DOCKET NUMBER: 028722-144  
TELEPHONE: 415-854-7400  
TELEFAX: 415-854-8275  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2061 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 16...900  
US-09-258-371-9

Alignment Scores:  
Pred. No.: 1.35e-71 Length: 2061  
Score: 728.00 Matches: 136  
Percent Similarity: 74.79% Conservative: 39  
Best Local Similarity: 58.12% Mismatches: 45  
Query Match: 49.16% Indels: 14  
Gaps: 2  
US-09-513-365A-1 (1-280) x US-09-258-371-9 (1-2061)

Qy 56 TyrGlnGlnThrLeuLysGluLeuAspValTyrGluLysTyrLysLysGluAsp 75  
Db 190 TGGAAACAGATCCCTGAAGAGCTAGACAGTGTACGAGCGCTTCAGTCGGAGACAG 249  
Qy 76 LeuAsnGlnLysLysArgLeuGlnLeuGlnArgAlaLeuLeuAsnSerGlnGlu 95  
Db 250 GGGCGGAGAGCGCGGATCTGCTGCTGTGCGCGCGCTGATCCGCGAGGAG 309  
Qy 96 LeuGlyAspGluLysIleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAla 115  
Db 310 CTGGCGGAGAGAGATCCAGATCGTGAGCCAGATGTTGGAGCTGTTGGAGAACCGCAG 369  
Qy 116 ArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGlu----- 130  
Db 370 CGGAGTGGACACGCCAGCTGAGCTGTTCGAGCGCGAGAGCTGGGCGACACAGTG 429  
Qy 131 -----SerGluArgAlaSerAspLysAlaLysMetAspSerSer 143  
Db 430 GGCAACAGCGCAAGTTGGCGGAGCGAGCCCGATGGCGATGGCGATGCGCGATCTGTAC 489  
Qy 144 GlnProGluArgSerArgArgProArgGlnArgThrSerGluSerArgAspLeu 163  
Db 490 AAGCCC-----AACAGCAAGCGCTCAGCGCGGAGCGCAACAGAGAACCGTGAGAAC 543  
Qy 164 CysHisMetAlaAsnGlyIleGluAspCysAspAspGlnProProLysGluLysSer 183  
Db 544 GCGTCCAGCAACACAGCCAGCGCGCGCTCGGCGCACACAGAGAGAGAGAGGCC 603  
Qy 184 LysSerAlaLysLysLysArgSerLysAlaLysGlnGluArgGluAlaSerProVal 203  
Db 604 AAGACCTCCAGAGAGAGAGCGCTCCAAAGCCCAAGCGGAGCGGAGCGGCGCTCCCGCTGCC 663

Qy 204 GluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnValSerTyrGly 223  
Db 664 GACCTCCCATCGACCCCAACGACGCTACTGTCTGTGCAACACGCTCTCTATGGG 723  
Qy 224 GluMetIleGlyCysAspAsnGluGlnCysProIleGluTrpPheHisPheSerCysVal 243  
Db 724 GAGATGATCGGCTGCGACACGAGAGTGCCTCATCGAGTGTTCACCTTCTCTGTCGCTG 783  
Qy 244 SerLeuThrTyrLysProLysGlyLysTrpTyrCysProLysCysArgGlyAspAsnGlu 263  
Db 784 GGGCTCAATCAATAACCCCAAGGCAAGTGTCTCCCAAGTGCCTGCGGAGAACGAG 843  
Qy 264 LysThrMetAspLysSerThrGluLysThrLysLysAspArg 277  
Db 844 AAGACCATGGACAAAGCCCTGGAGAAATCCAAAAAGAGAGG 885

RESULT 12  
US-08-751-230-9  
Sequence 9, Application US/08751230  
Patent No. 6117633  
GENERAL INFORMATION:  
APPLICANT: Garkavtsev, Igor  
TITLE OF INVENTION: DNA SEQUENCE ENCODING THE TUMOR  
SUPPRESSOR GENE INGI  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burns, Doane, Swecker & Mathis  
STREET: 699 Prince Street  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/751,230  
FILING DATE: 15-NOV-1996  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/569721  
FILING DATE: 08-DEC-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Mooi, Leslie A.  
REGISTRATION NUMBER: 37,047  
REFERENCE/DOCKET NUMBER: 028722-144  
TELEPHONE: 415-854-7400  
TELEFAX: 415-854-8275  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2061 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 16...900  
US-08-751-230-9

Alignment Scores:  
Pred. No.: 1.35e-71 Length: 2061  
Score: 728.00 Matches: 136  
Percent Similarity: 74.79% Conservative: 39  
Best Local Similarity: 58.12% Mismatches: 45  
Query Match: 49.16% Indels: 14  
Gaps: 2  
US-09-513-365A-1 (1-280) x US-08-751-230-9 (1-2061)

```
QY 56 TyrGlnGluThrLeuLysGluLeuAspValTyrGluLysTyrLysLysGluAspAsp 75
Db 190 TGGAAACAGATCTGAAGGAGTAGACGAGTGTACGAGCGTTCAGTCCGAGACAGAC 249
QY 76 LeuAsnGlnLysLysArgLeuGlnGlnLeuLeuLeuArgAlaLeuLeuAsnSerGlnGlu 95
Db 250 GGGCGCGCAAGCGCGGATGTGCACCTGTGTGACGCGCGTGTATCCGACGAGGAG 309
QY 96 LeuGlyAspGluLysIleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAla 115
Db 310 CTGGCGGCGAGAAGATCCAGATCGTGAGCCAGATGTTGAGTGTGTGAGAACCCGACG 369
QY 116 ArgGlnMetAlaAsnGlyIleGlnLysSerGlnCysPheGlnAspProAlaGlu 130
Db 370 CGGCGAGTGGACAGCCAGTGTTCGAGCGCGACGAGGAGTGGCGACACAGTG 429
QY 131 -----SerGluArgAlaSerAspLysAlaLysMetAspSerSer 143
Db 430 GGCACACGCGCAAGGTTGCGCGGACAGCCCAATGGCGATGGCGAGTCTGAC 489
QY 144 GlnProGluArgSerArgProArgProArgProArgProArgProArgProArgProVal 163
Db 490 AAGCCC-----AACAGCAAGCGCTCACGCGCGACGCAACACGAGAACCGTGAGAAC 543
QY 164 CysHisMetAlaAsnGlyIleGlnLysPheGlnCysAspGlnProLysGluLysSer 183
Db 544 GCGTCCAGCAACACGACGACGCGCGCTCGGCGACACACCCCAAGGAGAGAGGCC 603
QY 184 LysSerAlaLysLysLysArgSerLysAlaLysGlnGlnArgGluAlaSerProVal 203
Db 604 AAGACCTCCAAAGAGAGAGCGCTCCAGGCGCCAGGCGAGAGCGTCCCTGCC 663
QY 204 GluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnValSerTyrGly 223
Db 664 GACCTCCCATCGACCCCAACGACGACGCTACTGTGTGACACGAGTCTCCTATGGG 723
QY 224 GluMetIleGlyCysAspAsnGlnCysProIleGluTrpPheHisPheSerCysVal 243
Db 724 GAGATGATCGGTGCGACACGACGAGTGCCTCGGCGACGAGTGTTCCTCCTCGTGTG 783
QY 244 SerLeuThrTyrLysProLysGlyLysTyrTyrCysProLysCysArgGlyAspAsnGlu 263
Db 784 GGGCTCAATCAATAAACCCAAAGGCAAGTGTGTACTGTCCCAAGTCCCGGGGGAGAACGAG 843
QY 264 LysThrMetAspLysSerThrGluLysThrLysLysAspArg 277
Db 844 AAGACCATGGACAAAGCCCTGGAGAAATCCAAAAGAGAGG 885
```

## RESULT 13

US-09-499-082-9

Sequence 9, Application US/09499082

Patent No. 6143522

GENERAL INFORMATION:

APPLICANT: Helbing, Karen C.

APPLICANT: Riabowol, Karl

APPLICANT: Johnston, Randall N.

APPLICANT: Garkavtsev, Igor

TITLE OF INVENTION: METHODS OF MODULATING APOPTOSIS

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSEE: Burns, Doane, Swecker &amp; Mathis

STREET: 699 Prince Street

CITY: Alexandria

STATE: VA

COUNTRY: USA

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

```
APPLICATION NUMBER: US/09/499,082
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/828,158
FILING DATE: 27-MAR-1997
APPLICATION NUMBER: US 08/751230
FILING DATE: 15-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/569721
FILING DATE: 08-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Mooi, Leslie A.
REGISTRATION NUMBER: 37,047
REFERENCE/DOCKET NUMBER: 028722-148
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-854-7400
TELEFAX: 650-854-8275
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 2061 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 16...900
US-09-499-082-9
Alignment Scores:
Pred. No.: 1,35e-71 Length: 2061
Score: 728.00 Matches: 136
Percent Similarity: 74.79% Conservative: 39
Best Local Similarity: 58.12% Mismatches: 45
Query Match: 49.16% Indels: 14
DB: Gaps: 2
```

US-09-513-365A-1 (1-280) x US-09-499-082-9 (1-2061)

```
QY 56 TyrGlnGluThrLeuLysGluLeuAspValTyrGluLysTyrLysLysGluAspAsp 75
Db 190 TGGAAACAGATCTGAAGGAGTAGACGAGTGTACGAGCGTTCAGTCCGAGACAGAC 249
QY 76 LeuAsnGlnLysLysArgLeuGlnGlnLeuLeuLeuArgAlaLeuLeuAsnSerGlnGlu 95
Db 250 GGGCGCGCAAGCGCGGATGTGCACCTGTGTGACGCGCGTGTATCCGACGAGGAG 309
QY 96 LeuGlyAspGluLysIleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAla 115
Db 310 CTGGCGGCGAGAAGATCCAGATCGTGAGCCAGATGTTGAGTGTGTGAGAACCCGACG 369
QY 116 ArgGlnMetAlaAsnGlyIleGlnLysSerGlnCysPheGlnAspProAlaGlu 130
Db 370 CGGCGAGTGGACAGCCAGTGTTCGAGCGCGACGAGGAGTGGCGACACAGTG 429
QY 131 -----SerGluArgAlaSerAspLysAlaLysMetAspSerSer 143
Db 430 GGCACACGCGCAAGGTTGCGCGGACAGCCCAATGGCGATGGCGAGTCTGAC 489
QY 144 GlnProGluArgSerArgProArgProArgProArgProArgProArgProVal 163
Db 490 AAGCCC-----AACAGCAAGCGCTCACGCGCGACGCAACACGAGAACCGTGAGAAC 543
QY 164 CysHisMetAlaAsnGlyIleGlnLysPheGlnCysAspGlnProLysGluLysSer 183
Db 544 GCGTCCAGCAACACGACGACGCGCGCTCGGCGACACACCCCAAGGAGAGAGGCC 603
QY 184 LysSerAlaLysLysLysArgSerLysAlaLysGlnGlnArgGluAlaSerProVal 203
Db 604 AAGACCTCCAAAGAGAGAGCGCTCCAGGCGCCAGGCGAGAGCGTCCCTGCC 663
QY 204 GluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnValSerTyrGly 223
```

Db 664 GACCTCCCATCGACCCCAAGAGCCACCTACTGTGTGCAACCAAGGCTCTCTATGGG 723  
Qy 224 GluMetileGlyCysAspAsnGluInCysProileGluTrpPheHisPheSerCysVal 243  
Db 724 GAGATGATCGCTGCGCAACAGCAGAGTGCCTCATCGAGTGGTCCACTTCTCGTGCGTG 783  
Qy 244 SerLeuThrTyrLysProLysGlyLysTyrTrpCysProLysCysArgGlyAspAsnGlu 263  
Db 784 GGGCTCAATCATRAACCAAGGCAAGTGTCTGCTCCCAAGTCCCGGGGGAGAACGAG 843  
Qy 264 LysThrMetAspLysSerThrGluLysThrLysLysAspArg 277  
Db 844 AAGACCATGACAAAGCCCTGAGAAATCCAAAAGAGAGG 885

## RESULT 14

US-09-258-372-9  
; Sequence 9, Application US/09258372  
; Patent No. 6238918  
; GENERAL INFORMATION:  
; APPLICANT: Garkavtsev, Igor  
; APPLICANT: Riabowol, Karl  
; TITLE OF INVENTION: DNA SEQUENCE ENCODING THE TUMOR  
; TITLE OF INVENTION: SUPPRESSOR GENE INGI  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Burns, Doane, Swecker & Mathis  
; STREET: 699 Prince Street  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-1404  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/258,372  
; FILING DATE:

CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/751,230  
; FILING DATE:

ATTORNEY/AGENT INFORMATION:  
; NAME: Mool, Leslie A.  
; REGISTRATION NUMBER: 37,047  
; REFERENCE/DOCKET NUMBER: 028722-144

TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-854-7400  
; TELEFAX: 415-854-8275

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 2061 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 16..900

US-09-258-372-9

## Alignment Scores:

Pred. No.: 1.35e-71 Length: 2061  
Score: 728.00 Matches: 136  
Percent Similarity: 74.79% Conservative: 39  
Best Local Similarity: 58.12% Mismatches: 45  
Query Match: 49.16% Indels: 14  
DB: 4 Gaps: 2

US-09-513-365A-1 (1-280) x US-09-258-372-9 (1-2061)

Qy 56 TyrGlnGluThrLeuLysGluIleAspValTyrGluLysTyrLysLysGluAspAsp 75  
Db 190 TGGAAACAGATCTCTGAGGAGTACAGAGTGTACGCGCTTTCAGTCCGACAGACAG 249  
Qy 76 LeuAsnGlnLysLysArgLeuGlnLeuLeuGlnArgAlaLeuIleAsnSerGlnGlu 95  
Db 250 GGGCGCAGAGAGCGCGGATGCTGCTGCTGAGCGCGGCTGATCCGCGAGCAGAG 309  
Qy 96 LeuGlyAspGluLysIleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAla 115  
Db 310 CTGGCCACGAGAGATCCAGATCTGACGACATGTTGGAGCTGTGGAGAACCGCAGC 369  
Qy 116 ArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGlu----- 130  
Db 370 CGCAGGTGGACACGACGCTGCTGCGGCGCAGCAGGAGCTGGGCGACACAGTG 429  
Qy 131 -----SerGluArgAlaSerAspLysAlaLysMetAspSerSer 143  
Db 430 GGCAACAGCGCAAGGTTGGCGCGACAGGCCCAATGGCGATGGCGTACGCTCTGAC 489  
Qy 144 GlnProGluArgSerArgArgArgGlnArgThrSerGluSerArgAspLeu 163  
Db 490 AAGCCC-----AACACAGCGCTCAGCGCGGAGCGCAACACAGGACCGTGAGAAC 543  
Qy 164 CysHisMetAlaAsnGlyIleGluAspCysAspAspGlnProProLysGluLysSer 183  
Db 544 GCGTCCAGCAACACGACGACGCGCGCTCGGCGCACACCAAGGAGAAAGAGGCC 603  
Qy 184 LysSerAlaLysLysLysLysArgSerLysAlaLysGlnGluArgGluAlaSerProVal 203  
Db 604 AAGACCTCCAAGAAAGAGCGCTCCAAGCGGAGCGGAGCGGAGCGCTCCCTGCC 663  
Qy 204 GluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnValSerTyrGly 223  
Db 564 GACCTCCCATCGACCCCAACAGCAACCCACGACTGTCTGTGCAACACGAGTCTCTATGGG 723  
Qy 224 GluMetileGlyCysAspAsnGluInCysProileGluTrpPheHisPheSerCysVal 243  
Db 724 GAGATGATCGCTGCGCAACAGCAGAGTGCCTCATCGAGTGGTCCACTTCTCGTGCGTG 783  
Qy 244 SerLeuThrTyrLysProLysGlyLysTyrTrpCysProLysCysArgGlyAspAsnGlu 263  
Db 784 GGGCTCAATCATRAACCAAGGCAAGTGTGCTCCCAAGTCCCGGGGGAGAACGAG 843  
Qy 264 LysThrMetAspLysSerThrGluLysThrLysLysAspArg 277  
Db 844 AAGACCATGACAAAGCCCTGAGAAATCCAAAAGAGAGG 885

## RESULT 15

US-09-159-871-1  
; Sequence 1, Application US/09159871A  
; Patent No. 6420136  
; GENERAL INFORMATION:  
; APPLICANT: RIABOWOL, Karl T.  
; TITLE OF INVENTION: METHOD OF MODULATING P53 ACTIVITY  
; FILE REFERENCE: 028722-181  
; CURRENT APPLICATION NUMBER: US/09/159,871A  
; CURRENT FILING DATE: 1998-09-24  
; PRIOR APPLICATION NUMBER: US 60/060,138  
; PRIOR FILING DATE: 1997-09-26  
; PRIOR APPLICATION NUMBER: US 09/006,783  
; PRIOR FILING DATE: 1998-01-14  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 2061  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (16)..(897)  
US-09-159-871-1





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 20, 2003, 03:27:49 ; Search time 160 Seconds  
(without alignments)  
2567.998 Million cell updates/sec

Title: US-09-513-365A-1

Perfect score: 1481

Sequence: 1 MLGQQQQQLYSSAALLTGER.....DNEKTMDSKTEKTKDRRSR 280

Scoring table:

BLOSUM62,  
Xgapop 10.0 , Xgapext 0.5  
Xgapop 10.0 , Fgapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1042519 seqs, 733713590 residues

Total number of hits satisfying chosen parameters: 2085038

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO.spool/US09513365/runat\_16062003\_124745\_16556/app\_query.fasta\_1.455  
-DB=PublishedApplications\_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09513365 -CGN=1.1.80 -runat\_16062003\_124745\_16556  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications\_NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/PCTUS\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1481	100.0	840	12	US-10-115-899-6
2	1481	100.0	1078	12	US-10-115-899-7
3	853.5	57.6	873	9	US-09-968-653A-4
4	728	49.2	8487	10	US-09-764-877-3454

Sequence 6, Appli  
Sequence 7, Appli  
Sequence 4, Appli  
Sequence 3454, Ap

#### ALIGNMENTS

##### RESULT 1

US-10-115-899-6  
; Sequence 6, Application US/10115899  
; Patent No. US20020151025A1  
; GENERAL INFORMATION:  
; APPLICANT: Ostuka Pharmaceutical Co., Ltd.  
; TITLE OF INVENTION: Human TSC403 gene and human INGL1 gene  
; FILE REFERENCE: Q60193  
; CURRENT APPLICATION NUMBER: US/10/115.899  
; CURRENT FILING DATE: 2002-04-05  
; PRIOR APPLICATION NUMBER: 09/601,478  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: JP H10-134679  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: JP H10-73234  
; PRIOR FILING DATE: 1998-03-05  
; PRIOR APPLICATION NUMBER: JP H10-38133  
; PRIOR FILING DATE: 1998-02-03  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 840  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: human embryonic brain cDNA library  
US-10-115-899-6

5	722	48.8	2061	9	US-09-968-653A-2	Sequence 2, Appli
6	662	44.7	633	9	US-09-968-653A-6	Sequence 6, Appli
7	474	32.0	421	9	US-09-918-995-37033	Sequence 37033, A
8	359	24.2	1864	9	US-10-037-270-435	Sequence 435, App
9	314	21.2	451	9	US-09-854-133-146	Sequence 146, App
10	314	21.2	451	10	US-09-738-973-146	Sequence 146, App
11	277.5	18.7	813	9	US-09-938-842A-1178	Sequence 1178, Ap
12	263.5	17.8	993	10	US-09-801-368-257	Sequence 257, App
13	160	10.8	456	9	US-09-918-995-16047	Sequence 16047, A
14	130.5	8.8	7805	12	US-10-044-090-370	Sequence 370, App
15	129.5	8.7	915	10	US-09-764-864-471	Sequence 471, App
16	129.5	8.7	2320	9	US-10-175-523-142	Sequence 142, App
17	128.5	8.7	352	10	US-09-867-701-2399	Sequence 2399, Ap
18	128	8.6	3391	9	US-10-072-094-96	Sequence 96, Appl
19	128	8.6	4467	9	US-10-072-094-88	Sequence 88, Appl
20	127.5	8.6	1820	9	US-09-286-488-17	Sequence 17, Appl
21	127.5	8.6	1820	9	US-09-286-488-37	Sequence 37, Appl
22	127.5	8.6	1820	9	US-09-853-079-17	Sequence 17, Appl
23	127.5	8.6	1820	9	US-09-853-079-37	Sequence 37, Appl
24	127.5	8.6	1820	10	US-09-737-178-17	Sequence 17, Appl
25	127.5	8.6	1820	10	US-09-737-178-37	Sequence 37, Appl
26	127.5	8.6	2001	9	US-09-853-079-84	Sequence 84, Appl
27	127.5	8.6	2001	10	US-09-737-178-84	Sequence 84, Appl
28	127.5	8.6	2034	9	US-09-853-079-143	Sequence 143, App
29	127.5	8.6	2034	10	US-09-737-178-143	Sequence 210, App
30	127.5	8.6	2079	9	US-09-853-079-210	Sequence 86, Appl
31	127.5	8.6	3402	9	US-09-853-079-86	Sequence 86, Appl
32	127.5	8.6	3402	10	US-09-737-178-86	Sequence 86, Appl
33	125.5	8.5	2790	9	US-10-072-094-98	Sequence 98, Appl
34	125.5	8.5	3186	9	US-10-072-094-97	Sequence 97, Appl
35	125.5	8.5	3186	9	US-10-173-539-1	Sequence 1, Appli
36	125.5	8.5	3209	9	US-10-072-094-106	Sequence 106, App
37	125.5	8.5	3499	9	US-10-173-539-13	Sequence 3, Appli
38	125.5	8.5	3550	9	US-10-173-539-13	Sequence 13, Appl
39	125.5	8.5	4238	9	US-10-072-094-100	Sequence 100, App
40	125.5	8.5	7303	9	US-10-173-539-15	Sequence 15, Appl
41	125.5	8.5	7699	9	US-10-173-539-14	Sequence 14, Appl
42	125	8.4	5739	10	US-09-960-253-142	Sequence 142, App
43	125	8.4	6417	10	US-09-962-436-288	Sequence 288, App
44	125	8.4	6475	9	US-10-037-270-325	Sequence 325, App
45	123	8.3	1086	10	US-09-864-761-17120	Sequence 17120, A

## Alignment Scores:

Pred. No.: 4,4e-162 Length: 840  
Score: 1481.00 Matches: 280  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0

US-09-513-365A-1 (1-280) x US-10-115-899-6 (1-840)

```
QY 1 MetLeuGlyGlnGlnGlnGlnLeuTyrSerSerAlaAlaLeuLeuThrGlyGluArg 20
DB 1 ATGTAGGCGCAGCAGCAGCAGCAACTGTACTCGTCGGCGCGCTCTGCTACCGGGAGCGG 60
QY 21 SerArgLeuLeuThrCysTyrValGlnAspTyrLeuGluCysValGluSerLeuProHis 40
DB 61 AGCCGGCTGCTCACCCTGCTACGTGCGAGGACTACCTTGAGTGGCTGAGTCGCTGCCAC 120
QY 41 AspMetGlnArgAsnValSerValLeuArgGluLeuAspAsnLysTyrGlnGluThrLeu 60
DB 121 GACATGCAGAGCAAGCTGTCTGCTGCGAGAGCTGGACAACAATATCAAGAACGTTA 180
QY 61 LysGluIleAspAspValTyrGlnLysTyrLysLysGluAspLeuAsnGlnLysLys 80
DB 181 AAGGAAATTTGATGATGCTTACGAAAAATATAAGAAAGAGATGATTTAAACAGAGAA 240
QY 81 ArgLeuGlnGlnLeuLeuArgAlaLeuIleAsnSerGlnGluLeuGlyAspGluLys 100
DB 241 CGTCTACAGACCTTCTCCAGAGACACTTAATATAGTCAAGAATTTGGAGATGAAAAA 300
QY 101 IleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAlaArgGlnMetGluLeu 120
DB 301 ATACAGATGTTTACACAATGCTCGAATTTGGTGGAAATCGGCAAGCAAAATGAGTTA 360
QY 121 HisSerGlnCysPheGlnAspProAlaGluSerGluArgAlaSerAspLysAlaLysMet 140
DB 361 CACTCACAGTGTTCAGATCCTGCTGAAATGTAACGAGCTCGATATAAGCAAGATG 420
QY 141 AspSerSerGlnProGluArgSerArgArgProArgArgGlnArgThrSerGluSer 160
DB 421 GATTCAGCCACACAGAGAGATCTTCAAGAACACCCCGCAGCGGACAGTGAAGC 480
QY 161 ArgAspLeuCysHisMetAlaAsnGlyIleGluAspCysAspAspGlnProProLysGlu 180
DB 481 CGTGATTTATGTCACATGTCACAAATGGGATTTGAAGACTGTGATGATCAGCCACTAAAG 540
QY 181 LysLysSerLysSerAlaLysLysLysLysArgSerLysAlaLysGlnGluArgGluAla 200
DB 541 AAGAAATCCAAGTCAGCAAGAAAGAAAGAACCGCTCCAGGCCAAGCAGGAAAGGAGCT 600
QY 201 SerProValGluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnVal 220
DB 601 TCACCTGTGTGAGTTGTCATAGATCCTATGAACTACATCTGCTTATGCAACCAAGTG 660
QY 221 SerTyrGlyGluMetIleGlyCysAspAsnGluGlnCysProIleGluTrpPheHisPhe 240
DB 661 TCTTATGGGGAGATGATAGGATGTGACATGACATGAACAGTGTCCAAATGAATGTTTCACTT 720
QY 241 SerCysValSerLeuThrTyrLysProLysGlyLysTyrTrpTyrCysProLysCysArgGly 260
DB 721 TCATGTGTTTCACTTACCTTATAACCAAGGGGAATGTTATGCCAAAGTGCAGGGGA 780
QY 261 AspAsnGluLysThrMetAspLysSerThrGluLysThrLysLysAspArgArgSerArg 280
DB 781 GATAATGAGAAACAATGTCACAAAGTACTGAAAGACAAAAAGATAGATAGATCGAGG 840
```

## RESULT 2

US-10-115-899-7

; Sequence 7, Application US/10115899  
; Patent No. US20020151025A1  
; GENERAL INFORMATION:  
; APPLICANT: Ostuka Pharmaceutical Co., Ltd.  
; TITLE OF INVENTION: Human TSC403 gene and human ING1L gene

; FILE REFERENCE: Q60193  
; CURRENT APPLICATION NUMBER: US/10/1115,899  
; CURRENT FILING DATE: 2002-04-05  
; PRIOR APPLICATION NUMBER: 09/601,478  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: JP H10-134679  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: JP H10-73234  
; PRIOR FILING DATE: 1998-03-05  
; PRIOR APPLICATION NUMBER: JP H10-38133  
; PRIOR FILING DATE: 1998-02-03  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 1078  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: human embryonic brain cDNA library  
; NAME/KEY: CDS  
; LOCATION: (92)..(931)  
; US-10-115-899-7

## Alignment Scores:

Pred. No.: 6,29e-162 Length: 1078  
Score: 1481.00 Matches: 280  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0

US-09-513-365A-1 (1-280) x US-10-115-899-7 (1-1078)

```
QY 1 MetLeuGlyGlnGlnGlnGlnLeuTyrSerSerAlaAlaLeuLeuThrGlyGluArg 20
DB 92 ATGTAGGCGCAGCAGCAGCAGCAACTGTACTCGTCGGCGCGCTCTGCTACCGGGAGCGG 151
QY 21 SerArgLeuLeuThrCysTyrValGlnAspTyrLeuGluCysValGluSerLeuProHis 40
DB 152 AGCCGGCTGCTCACCCTGCTACGTGCGAGGACTACCTTGAGTGGCTGAGTCGCTGCCAC 211
QY 41 AspMetGlnArgAsnValSerValLeuArgGluLeuAspAsnLysTyrGlnGluThrLeu 60
DB 212 GACATGCAGAGCAAGCTGTCTGCTGCGAGAGCTGGACAACAATATCAAGAACGTTA 271
QY 61 LysGluIleAspAspValTyrGlnLysTyrLysLysGluAspAspLeuAsnGlnLysLys 80
DB 272 AAGGAATTTGATGATGCTTACGAAAAATATAAGAAAGAGATGATTTAAACCAAGAGAA 331
QY 81 ArgLeuGlnGlnLeuLeuGlnArgAlaLeuIleAsnSerGlnGluLeuGlyAspGluLys 100
DB 332 CGTCTACAGCAGCTTCTCCAGAGACACTTAATTAATAGTCAAGAATTTGGAGATGAAAAA 391
QY 101 IleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAlaArgGlnMetGluLeu 120
DB 392 ATACAGATGTTTACACAATGCTCGAATTTGGTGGAAATCGGCGCAAGCAAAATGAGTTA 451
QY 121 HisSerGlnCysPheGlnAspProAlaGluSerGluArgAlaSerAspLysAlaLysMet 140
DB 452 CACTCACAGTGTTCAGATCCTGCTGAAAGTGAACGAGCTCGATATAAGCAAGAGTG 511
QY 141 AspSerSerGlnProGluArgSerArgArgProArgArgGlnArgThrSerGluSer 160
DB 512 GATTCAGCAACACAGAGAGATCTTCAAGAGACCCCGCAGCGGACAGTGAAGC 571
QY 161 ArgAspLeuCysHisMetAlaAsnGlyIleGluAspCysAspAsnGlnProProLysGlu 180
DB 572 CGTGATTTATGTCACATGCGCAATGGGATTTGAAGACTGTGATGATCAGCGCCCTAAAGAA 631
QY 181 LysLysSerLysSerAlaLysLysLysLysArgSerLysAlaLysGlnGluArgGluAla 200
DB 632 AAGAAATCCAAGTCAGCAAGAAAGAAAGAACCGCTCCAGGCCAAGCAGGAAAGGAGCT 691
```

Qy 201 SerProValGluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnVal 220  
 Db 692 TCACCTGTGAGTTGCAATAGATCTTAATGAACCTACATACCTGTTATGCAACCAAGTG 751  
 Qy 221 SerTyrGluMetIleGlyCysAspAsnGluGlnCysProIleGluTrpPheHisPhe 240  
 Db 752 TCTTATGGGAGATGATAGGATGACAAATGAACAGTGTCCAAATGAATGTTTCACTTT 811  
 Qy 241 SerCysValSerLeuThrTyrLysProLysGlyLysTrpTyrCysProLysCysAspArgGly 260  
 Db 812 TCATGTGTTCACTTACCTATACCAAGGGAATGGTATGCCCCAAAGTGCGAGGGA 871  
 Qy 261 AspAsnGluLysThrMetAspLysSerThrGluLysThrLysLysAspArgSerArg 280  
 Db 872 GATAATGAGAAACAAATGGACAAAGTACTGAAAGACAAAGAGGATAGAGATCGAGG 931

RESULT 3

US-09-968-653A-4  
 ; Sequence 4, Application US/09968653A  
 ; Publication No. US20030073084A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gudkov, Andrey V  
 ; Garbavstev, Igor  
 ; Riabowol, Karl  
 ; TITLE OF INVENTION: p33/ING1 as a Mediator of p53 Signaling  
 ; NUMBER OF SEQUENCES: 7  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
 ; STREET: 300 South Wacker Drive  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: USA  
 ; ZIP: 60606  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/968,653A  
 ; FILING DATE: 01-Oct-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/006,783A  
 ; FILING DATE: 15-JAN-1998  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: No. US20030073084A1nan, Kevin E  
 ; REGISTRATION NUMBER: 35,303  
 ; REFERENCE/DOCKET NUMBER: 97,837  
 ; TELEPHONE: 312-913-0001  
 ; TELEFAX: 312-913-0002  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 873 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 7..813  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-968-653A-4  
 ; Sequence 4, Application US/09968653A  
 ; Publication No. US20030073084A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gudkov, Andrey V  
 ; Garbavstev, Igor  
 ; Riabowol, Karl  
 ; TITLE OF INVENTION: p33/ING1 as a Mediator of p53 Signaling  
 ; NUMBER OF SEQUENCES: 7  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
 ; STREET: 300 South Wacker Drive  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: USA  
 ; ZIP: 60606  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/968,653A  
 ; FILING DATE: 01-Oct-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/006,783A  
 ; FILING DATE: 15-JAN-1998  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: No. US20030073084A1nan, Kevin E  
 ; REGISTRATION NUMBER: 35,303  
 ; REFERENCE/DOCKET NUMBER: 97,837  
 ; TELEPHONE: 312-913-0001  
 ; TELEFAX: 312-913-0002  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 873 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 7..813  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-968-653A-4  
 ; Sequence 4, Application US/09968653A  
 ; Publication No. US20030073084A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gudkov, Andrey V  
 ; Garbavstev, Igor  
 ; Riabowol, Karl  
 ; TITLE OF INVENTION: p33/ING1 as a Mediator of p53 Signaling  
 ; NUMBER OF SEQUENCES: 7  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
 ; STREET: 300 South Wacker Drive  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: USA  
 ; ZIP: 60606  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/968,653A  
 ; FILING DATE: 01-Oct-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/006,783A  
 ; FILING DATE: 15-JAN-1998  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: No. US20030073084A1nan, Kevin E  
 ; REGISTRATION NUMBER: 35,303  
 ; REFERENCE/DOCKET NUMBER: 97,837  
 ; TELEPHONE: 312-913-0001  
 ; TELEFAX: 312-913-0002  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 873 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 7..813  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-968-653A-4  
 ; Sequence 4, Application US/09968653A  
 ; Publication No. US20030073084A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gudkov, Andrey V  
 ; Garbavstev, Igor  
 ; Riabowol, Karl  
 ; TITLE OF INVENTION: p33/ING1 as a Mediator of p53 Signaling  
 ; NUMBER OF SEQUENCES: 7  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
 ; STREET: 300 South Wacker Drive  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: USA  
 ; ZIP: 60606  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/968,653A  
 ; FILING DATE: 01-Oct-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/006,783A  
 ; FILING DATE: 15-JAN-1998  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: No. US20030073084A1nan, Kevin E  
 ; REGISTRATION NUMBER: 35,303  
 ; REFERENCE/DOCKET NUMBER: 97,837  
 ; TELEPHONE: 312-913-0001  
 ; TELEFAX: 312-913-0002  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 873 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 7..813  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-968-653A-4  
 ; Sequence 4, Application US/09968653A  
 ; Publication No. US20030073084A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gudkov, Andrey V  
 ; Garbavstev, Igor  
 ; Riabowol, Karl  
 ; TITLE OF INVENTION: p33/ING1 as a Mediator of p53 Signaling  
 ; NUMBER OF SEQUENCES: 7  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
 ; STREET: 300 South Wacker Drive  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: USA  
 ; ZIP: 60606  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/968,653A  
 ; FILING DATE: 01-Oct-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/006,783A  
 ; FILING DATE: 15-JAN-1998  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: No. US20030073084A1nan, Kevin E  
 ; REGISTRATION NUMBER: 35,303  
 ; REFERENCE/DOCKET NUMBER: 97,837  
 ; TELEPHONE: 312-913-0001  
 ; TELEFAX: 312-913-0002  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 873 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 7..813  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-968-653A-4  
 ; Sequence 4, Application US/09968653A  
 ; Publication No. US20030073084A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gudkov, Andrey V  
 ; Garbavstev, Igor  
 ; Riabowol, Karl  
 ; TITLE OF INVENTION: p33/ING1 as a Mediator of p53 Signaling  
 ; NUMBER OF SEQUENCES: 7  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
 ; STREET: 300 South Wacker Drive  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: USA  
 ; ZIP: 60606  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/968,653A  
 ; FILING DATE: 01-Oct-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/006,783A  
 ; FILING DATE: 15-JAN-1998  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: No. US20030073084A1nan, Kevin E  
 ; REGISTRATION NUMBER: 35,303  
 ; REFERENCE/DOCKET NUMBER: 97,837  
 ; TELEPHONE: 312-913-0001  
 ; TELEFAX: 312-913-0002  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 873 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 7..813  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-968-653A-4  
 ; Sequence 4, Application US/09968653A  
 ; Publication No. US20030073084A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gudkov, Andrey V  
 ; Garbavstev, Igor  
 ; Riabowol, Karl  
 ; TITLE OF INVENTION: p33/ING1 as a Mediator of p53 Signaling  
 ; NUMBER OF SEQUENCES: 7  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
 ; STREET: 300 South Wacker Drive  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: USA  
 ; ZIP: 60606  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/968,653A  
 ; FILING DATE: 01-Oct-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/006,783A  
 ; FILING DATE: 15-JAN-1998  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: No. US20030073084A1nan, Kevin E  
 ; REGISTRATION NUMBER: 35,303  
 ; REFERENCE/DOCKET NUMBER: 97,837  
 ; TELEPHONE: 312-913-0001  
 ; TELEFAX: 312-913-0002  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 873 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 7..813  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-968-653A-4  
 ; Sequence 4, Application US/09968653A  
 ; Publication No. US20030073084A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gudkov, Andrey V  
 ; Garbavstev, Igor  
 ; Riabowol, Karl  
 ; TITLE OF INVENTION: p33/ING1 as a Mediator of p53 Signaling  
 ; NUMBER OF SEQUENCES: 7  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
 ; STREET: 300 South Wacker Drive  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: USA  
 ; ZIP: 60606  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/968,653A  
 ; FILING DATE: 01-Oct-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/006,783A  
 ; FILING DATE: 15-JAN-1998  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: No. US20030073084A1nan, Kevin E  
 ; REGISTRATION NUMBER: 35,303  
 ; REFERENCE/DOCKET NUMBER: 97,837  
 ; TELEPHONE: 312-913-0001  
 ; TELEFAX: 312-913-0002  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 873 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 7..813  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-968-653A-4  
 ; Sequence 4, Application US/09968653A  
 ; Publication No. US20030073084A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gudkov, Andrey V  
 ; Garbavstev, Igor  
 ; Riabowol, Karl  
 ; TITLE OF INVENTION: p33/ING1 as a Mediator of p53 Signaling  
 ; NUMBER OF SEQUENCES: 7  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
 ; STREET: 300 South Wacker Drive  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: USA  
 ; ZIP: 60606  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/968,653A  
 ; FILING DATE: 01-Oct-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/006,783A  
 ; FILING DATE: 15-JAN-1998  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: No. US20030073084A1nan, Kevin E  
 ; REGISTRATION NUMBER: 35,303  
 ; REFERENCE/DOCKET NUMBER: 97,837  
 ; TELEPHONE: 312-913-0001  
 ; TELEFAX: 312-913-0002  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 873 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 7..813  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-968-653A-4  
 ; Sequence 4, Application US/09968653A  
 ; Publication No. US20030073084A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gudkov, Andrey V  
 ; Garbavstev, Igor  
 ; Riabowol, Karl  
 ; TITLE OF INVENTION: p33/ING1 as a Mediator of p53 Signaling  
 ; NUMBER OF SEQUENCES: 7  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
 ; STREET: 300 South Wacker Drive  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: USA  
 ; ZIP: 60606  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/968,653A  
 ; FILING DATE: 01-Oct-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/006,783A  
 ; FILING DATE: 15-JAN-1998  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: No. US20030073084A1nan, Kevin E  
 ; REGISTRATION NUMBER: 35,303  
 ; REFERENCE/DOCKET NUMBER: 97,837  
 ; TELEPHONE: 312-913-0001  
 ; TELEFAX: 312-913-0002  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 873 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 7..813  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-968-653A-4  
 ; Sequence 4, Application US/09968653A  
 ; Publication No. US20030073084A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gudkov, Andrey V  
 ; Garbavstev, Igor  
 ; Riabowol, Karl  
 ; TITLE OF INVENTION: p33/ING1 as a Mediator of p53 Signaling  
 ; NUMBER OF SEQUENCES: 7  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
 ; STREET: 300 South Wacker Drive  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: USA  
 ; ZIP: 60606  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/968,653A  
 ; FILING DATE: 01-Oct-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/006,783A  
 ; FILING DATE: 15-JAN-1998  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: No. US20030073084A1nan, Kevin E  
 ; REGISTRATION NUMBER: 35,303  
 ; REFERENCE/DOCKET NUMBER: 97,837  
 ; TELEPHONE: 312-913-0001  
 ; TELEFAX: 312-913-0002  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 873 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 7..813  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-968-653A-4  
 ; Sequence 4, Application US/09968653A  
 ; Publication No. US20030073084A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gudkov, Andrey V  
 ; Garbavstev, Igor  
 ; Riabowol, Karl  
 ; TITLE OF INVENTION: p33/ING1 as a Mediator of p53 Signaling  
 ; NUMBER OF SEQUENCES: 7  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
 ; STREET: 300 South Wacker Drive  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: USA  
 ; ZIP: 60606  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/968,653A  
 ; FILING DATE: 01-Oct-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/006,783A  
 ; FILING DATE: 15-JAN-1998  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: No. US20030073084A1nan, Kevin E  
 ; REGISTRATION NUMBER: 35,303  
 ; REFERENCE/DOCKET NUMBER: 97,837  
 ; TELEPHONE: 312-913-0001  
 ; TELEFAX: 312-913-0002  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 873 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 7..813  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-968-653A-4  
 ; Sequence 4, Application US/09968653A  
 ; Publication No. US20030073084A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gudkov, Andrey V  
 ; Garbavstev, Igor  
 ; Riabowol, Karl  
 ; TITLE OF INVENTION: p33/ING1 as a Mediator of p53 Signaling  
 ; NUMBER OF SEQUENCES: 7  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
 ; STREET: 300 South Wacker Drive  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: USA  
 ; ZIP: 60606  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/968,653A  
 ; FILING DATE: 01-Oct-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/006,783A  
 ; FILING DATE: 15-JAN-1998  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: No. US20030073084A1nan, Kevin E  
 ; REGISTRATION NUMBER: 35,303  
 ; REFERENCE/DOCKET NUMBER: 97,837  
 ; TELEPHONE: 312-913-0001  
 ; TELEFAX: 312-913-0002  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 873 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 7..813  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-968-653A-4  
 ; Sequence 4, Application US/09968653A  
 ; Publication No. US20030073084A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gudkov, Andrey V  
 ; Garbavstev, Igor  
 ; Riabowol, Karl  
 ; TITLE OF INVENTION: p33/ING1 as a Mediator of p53 Signaling  
 ; NUMBER OF SEQUENCES: 7  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
 ; STREET: 300 South Wacker Drive  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: USA  
 ; ZIP: 60606  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/968,653A  
 ; FILING DATE: 01-Oct-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/006,783A  
 ; FILING DATE: 15-JAN-1998  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: No. US20030073084A1nan, Kevin E  
 ; REGISTRATION NUMBER: 35,303  
 ; REFERENCE/DOCKET NUMBER: 97,837  
 ; TELEPHONE: 312-913-0001  
 ; TELEFAX: 312-913-0002  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 873 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 7..813  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-513-365A-1 (1-280) x US-09-968-653A-4 (1-873)  
 Qy 18 GlyGluArgSerArgLeuLeuThrCysTyrValGlnAspTyrLeuGluCysValGluSer 37  
 Db 25 GGGAGGAGGAGTCCACCTGGTGAAC---TATGGAGGAGTACCTGGAGTCCATCGAGTCC 81  
 Qy 38 LeuProHisAspMetGlnArgAsnValSerValLeuArgGluLeuAspAsnLysTyrGln 57  
 Db 82 CTGCTTTCGACTTGCAGAGAAATGTCGCTGATCGGGAGATCGAGCGCGAATACCAA 141  
 Qy 58 GluThrLeuLysGluIleAspValTyrGluLysTyrLysLysGluAspAsnLeu 77  
 Db 142 GAGATCTCTGAGGAGTACGAGTGTCTACGAGCGCTTCAGTCGCGACACAGACGGCGC 201  
 Qy 78 GlnLysLysArgLeuGlnGlnLeuLeuGlnArgAlaLeuIleAsnSerGlnGluLeu 97  
 Db 202 CAGAGCGGGGATGCTGCTGTCGACCGCGCGCTGATCCGCGAGCAGGAGCTGGGC 261  
 Qy 98 AspGluLysIleGlnIleValThrGlnMetLeuGluLeuValGluAsnArgAlaArgGln 117  
 Db 262 GACGAGAAGATCCAGATCGTCGACGAGATGTCGAGCTGTGGAGTGTGGAGAACCGCAG 321  
 Qy 118 MetGluLeuHisSerGlnCysPheGlnAspProAlaGlu----- 130  
 Db 322 GTGACACGCCACGTGGAGCTGTCGAGCGCGAGGAGTGGCGGACACAGTGGGCAAC 381  
 Qy 131 -----SerGluArgAlaSerAspLysAlaLysMetAspSerSerGlnPro 145  
 Db 382 AGCGGCAAGTTGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 441  
 Qy 146 GluArgSerArgArgProArgGlnArgThrSerGluSerArgAspLeuCysHis 165  
 Db 442 -----AACGCAAGCGCTCACGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 495  
 Qy 166 MetAlaAsnGlyIleGluAspCysAspAspGlnProProLysGluLysSerLysSer 185  
 Db 496 AGCAACACCGACCGACCGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 555  
 Qy 186 AlaLysLysLysLysArgSerLysAlaLysGlnGluArgGluAlaSerProValGluPhe 205  
 Db 556 TCCAGAGAGAAGAGCGCTCCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 615  
 Qy 206 AlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnValSerTyrGlyGluMet 225  
 Db 616 CCCATCGACCCCAACGAAACCCACGCTACTGTCTGTGCAACAGGCTCTCTATGGGAGATG 675  
 Qy 226 IleGlyCysAspAsnGluGlnCysProIleGluTrpPheHisPheSerCysValSerLeu 245  
 Db 676 ATCGGCTGCGACAACGAGAGTGCCTCCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 735  
 Qy 246 ThrTyrLysProLysGlyLysTrpTyrCysProLysCysArgGlyAspAsnGluLysThr 265  
 Db 736 AATCATAAACCCAGGCGCAAGTGTACTGTCTCCCAAGTCCCGGGGGGAGACAGAGAGACC 795  
 Qy 266 MetAspLysSerThrGluLysThrLysLysAspArg 277  
 Db 796 ATGGACAAAGCGCTGGAGAAATCCAAAAAGAGAGG 831

RESULT 4

US-09-764-877-3454  
 ; Sequence 3454, Application US/09764877  
 ; Patent No. US20020147140A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 ; FILE REFERENCE: PC005  
 ; CURRENT APPLICATION NUMBER: US/09/764,877  
 ; CURRENT FILING DATE: 2001-01-17  
 ; Prior application data removed - refer to PALM or file wrapper  
 ; NUMBER OF SEQ ID NOS: 4031  
 ; SOFTWARE: PatentIn ver. 2.0  
 ; SEQ ID NO 3454

Alignment Scores:  
 Pred. No.: 2,15e-89 Length: 873  
 Score: 853.50 Matches: 161  
 Percent Similarity: 75.74% Conservative: 45  
 Best Local Similarity: 59.19% Mismatches: 51  
 Query Match: 57.63% Indels: 15  
 DB: 9 Gaps: 3



QY 164 CysHisMetAlaAsnGlyIleGluAspCysAspGlnProProLysGluLysLysSer 183  
Db 544 CGCTCCAGCAACACACGACGAGCGGCGCTCGGCACACCCCAAGAGAGAGGCC 603  
QY 184 LysSerAlaLysLysLysArgSerLysAlaLysGlnGluArgGluAlaSerProVal 203  
Db 604 AAGACCTCAAGAAGAGAGCGCTCCAGGCGCAAGCGGAGGAGCGCTCCCTGCC 663  
QY 204 GluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnGlnValSerTyrGly 223  
Db 664 GACCTCCCATCGACCCCAAGCAACGCTACTGTGTGCAACGAGGTCTCTATGG 723  
QY 224 GluMetIleGlyCysAspAsnGlnCysProIleGluTrpPheHisPheSerCysVal 243  
Db 724 GAGATGATCGCTGCGACACGAGAGTGCCTCCATCGAGTGTTCCTCCTCGGTG 783  
QY 244 SerLeuThrTyrLysProLysGlyLysTrpTyrCysProLysCysArgGlyAspAsnGlu 263  
Db 784 GGCTCAATCAATAACCAAGCAAGTGTGTACTGTCCCAAGTGCCTGGGGGAGACGAG 843  
QY 264 LysThrMetAspLysSerThrGluLysThrLysLysAspArg 277  
Db 844 AAGACCATGACAAAGCCCTGGAGAAATCCAAAAAGAGAGG 885

RESULT 6

US-09-968-653A-6  
; Sequence 6, Application US/09968653A  
; Publication No. US20030073084A1  
; GENERAL INFORMATION:  
; APPLICANT: Gudkov, Andrey V  
; Rjabowol, Karl  
; TITLE OF INVENTION: p33/TING1 as a Mediator of p53 Signaling  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
; STREET: 300 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/968,653A  
; FILING DATE: 01-Oct-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/006,783A  
; FILING DATE: 15-JAN-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. US20030073084Alban, Kevin E  
; REGISTRATION NUMBER: 35,303  
; REFERENCE/DOCKET NUMBER: 97,837  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-913-0001  
; TELEFAX: 312-913-0002  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 633 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..630  
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-09-968-653A-6  
Alignment Scores:  
Pred. No.: 2,04e-67 Length: 633  
Score: 662.00 Matches: 125  
Percent Similarity: 75.49% Conservative: 29  
Best Local Similarity: 61.27% Mismatches: 36  
Query Match: 44.70% Indels: 14  
DB: 9 Gaps: 2  
US-09-513-365A-1 (1-280) x US-09-968-653A-6 (1-633)  
QY 86 LeuGlnArgAlaLeuIleAsnSerGlnGluLeuGlyAspGluLysIleGlnIleValThr 105  
Db 13 GTCAGCGCGCGCTGATCCGACGCCAGGAGCTGGCGACGAGAAGATCCAGATCGTGAGC 72  
QY 106 GlnMetLeuGluLeuValGluAsnArgAlaArgGlnMetGluLeuHisSerGlnCysPhe 125  
Db 73 CAGATGGTGGAGCTGTGGAGAACCGCACCGCGAGTGGACAGCCAGCTGCTTTC 132  
QY 126 GlnAspProAlaGlu-----SerGluArg 133  
Db 133 GAGCGCAGCAGGAGCTGGCGGACAGTGGGCAACAGCGCGAAGGTGGCGGCGGACAGG 192  
QY 134 AlaSerAspLysAlaLysMetAspSerSerGlnProGluArgSerSerArgArgProArg 153  
Db 193 CCCAATGGCGATCGGCTAGCGCAGTCTGACAAGCCC-----AACAGCAAGCGCTCACGG 246  
QY 154 ArgGlnArgThrSerGluSerArgAspLeuCysHisMetAlaAsnGlyIleGluAspCys 173  
Db 247 CGGCAGCGCAACACGAGAACCGCTGAGACCGCTCCAGCAACACCGACGACGACGCGC 306  
QY 174 AspAspGlnProProLysGluLysLysSerLysSerAlaLysLysLysLysLysLys 193  
Db 307 GCCTCGGGCACACCCAGGAGAGAGCGCCACAGCTCCCAAGAGAAAGAGCGCTCCAG 366  
QY 194 AlaLysGlnGluArgGluAlaSerProValIleGluPheAlaIleAspProAsnGluProThr 213  
Db 367 GCCAAGCGGAGCGAGAGCGCTCCCTCGCGGAGTATGATGCGGAGATGATGCGGAGAGTGC 426  
QY 214 TyrCysLeuCysAsnGlnValSerTyrGlyGluMetIleGlyCysAspAsnGluCys 233  
Db 427 TACTGTCTGTGCAACCGAGTCTCTATGGGAGATGATGCGGAGTGCAGACGAGAGTGC 486  
QY 234 ProIleGluTrpPheHisPheSerCysValSerLeuThrTyrLysProLysGlyLysTrp 253  
Db 487 CCCATCGAGTGGTTCACCTTCTCGTGGCTGGGCTCAATCATAAACCCAGGCGCAAGTGG 546  
QY 254 TyrCysProLysCysArgGlyAspAsnGluLysThrMetAspLysSerThrGluLysThr 273  
Db 547 TACTGTCCCAAGTCCCGGGGAGAGACGAGAGACCATGGACAAAGCCCTGGAGAAATCC 606  
QY 274 LysLysAspArg 277  
Db 607 AAAAAAGAGAGG 618

RESULT 7

US-09-918-995-37033  
; Sequence 37033, Application US/09918995  
; Publication No. US20030073623A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
; FILE OF INVENTION: FROM VARIOUS CDNA LIBRARIES  
; FILE REFERENCE: 20411-756  
; CURRENT APPLICATION NUMBER: US/09/918,995  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US/09/235,076  
; PRIOR FILING DATE: 1999-01-20  
; NUMBER OF SEQ ID NOS: 38054  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 37033  
; LENGTH: 421

1864  
102  
53  
100  
198  
11

US-09-513-365A-1 (1-280) x US-10-037-270-435 (1-1864)

[illegible]

Db	443	TTGAGAAAGCTGGATCAGGAAGCTGGCTTAAGTTTAAAATGGAGCTGGAGCTGATATATGCT	502
Qy	134	-----	134
Db	503	GGAATTCAGAAATATTAGACAGGCGATCTTTGGAAATTACACACTCTTCACAGCCAGTG	562
Qy	134	-----	134
Db	563	AACAATCACCATGCTCATTCACATACTCCAGTGGAAAAAGGAAATATAATCCAACTTCT	622
Qy	134	-----	134
Db	623	CACCATACGACACACAGATCATATTTCCTGAAAGAAATTTAAATCTGAAGCTCTTCTATCC	682
Qy	135	-----SerAspLysAlaLys-----	139
Db	683	ACCTTACGTCAGATGCCTCTAAGGAAAAATACACTAGTTGTCGAAATAATAATCCACA	742
Qy	140	-----MetAspSerSerGlnPro-----	145
Db	743	GCCTCTTTCTAACAATGCCTACAATGTGAATTTCTCCCAACCTCTGGGTCGTATAACATT	802
Qy	145	-----	145
Db	803	GGCTCGTTATCTTCAGGAAGCTGGTGACGGGGCAGTTACCATGGCAGCTGCTCAAGCAGTT	862
Qy	146	GluArgSerSerArgArgProArgGlnArgThrSerGlu-----	159
Db	863	CAGGCTACAGCTCAGATCAAGGAGGAGGACGAGACATCAAGTTTAAAGCCAGTTATGAA	922
Qy	159	-----	159
Db	923	GCATTTAAGAATAATGACTTTTTCAGTTGGGAAAAAGAAATTTTCAATGGCCAGGAAACAGTT	982



```

Db 137 GACGGGGCAGAGCGGGATGCTGCTGTCAGCGCGCTGATCCGAC-CAG 195
QY 95 GluLeuGlyAspGluLysIleGlnIleValThrGlnMetLeuValGluAsnArg 114
Db 196 GAGCTGGGCGACGAGAGATCCAGATCGTGACCATGCTGGTGAGACCCG 255
QY 115 AlaArgGlnMetLeuLeuHisSerGlnCysPheGlnAspProAlaGlu---SerGluArg 133
Db 256 ACGGCGCAGTGGACGCCAGCGTGGAGCTGCTCGAGCGCAGCAGCGTGGCGACACA 315
QY 134 AlaSerAspLysAlaLysMetAspSerSerGlnPro----- 145
Db 316 GCGGCAACAGCGCGCAAGCTGGCGCGACAGGCCCAAGCGCGCAGCGAGGCT. 375
QY 146 ---GluArgSerSerArgProArgAlaGlnThrSerGlnSerArgasp 162
Db 376 GACAAGCCCAACAGCAAGCGCTCACGGCGCAGCGCAACAGCAAGCGGTGAG 429

RESULT 11
US-09-938-842A-1178
; Sequence 1178, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1178
; LENGTH: 813
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1178

Alignment Scores:
Pred. No.: 9,79e-23 Length: 813
Score: 277.50 Matches: 77
Percent Similarity: 45.35% Conservativity: 40
Best Local Similarity: 29.84% Mismatches: 81
Query Match: 18.74% Indels: 61
DB: 9 Gaps: 9

US-09-513-365A-1 (1-280) x US-09-938-842A-1178 (1-813)
QY 27 TyrValGlnAspTyrLeuGluCysValGluSerLeuProHisAspMetGlnArgAsnVal 46
Db 25 TATGTTGATGATTACCTTCTGATGATGCAAGCACTTCCCTGCAGAGCTACAAAGATTACTA 84
QY 47 SerValLeuArgLysLeuAsp----- 53
Db 85 AACACAGTTCGGCAAGCTAGCAGAGATCTCAATGGCTTCTTCACAGAGTTCTTAAAG 144
QY 54 -----AsnLysTyrGlnGluThrLeuLysGluIleAspAspValTyrGluLysTyr 70
Db 145 GGTATGCTGAATCATTAATCAATATGTTGCTGCTGATGAAGAGAAACGATTGAGAAATG 204
QY 71 LysLysGluAspAspLeuAsnGlnLysLysArgLeuGlnGlnLeuLeuGlnArgAlaLeu 90
Db 205 CGTAAAGAGATTGAGTCTAGTCAGGAA----- 231
QY 91 IleAsnSerGlnGluLeuGlyAspGluLysIleGlnIleValThrGlnMetLeuGluLeu 110

```

```

Db 232 ---AATGCGTTAAGTTTGTGTCACGAGAAGGTTTATTGTCGCCGACAAGGTATGATCTT 288
QY 111 ValGluAsnArgAlaArgGlnMetGluLeuHisSerGlnCysPheGlnAspProAlaGlu 130
Db 289 ATAGATGATCATGTAAACGACGTTGATGAAGATCTGAATAATTTT-----GCAGAA 339
QY 131 SerGluArgAlaSerAspLysAlaLysMetAspSerSerGln----- 144
Db 340 GATTTAAAGCAGAGGGAATAATTCACACAGAGAGCCCTCTCTTCTCTCCACTACCT 399
QY 145 -----ProGluArgSerSerArgArg-ProArgArgGlnArgThrSerGluSerArgAs 162
Db 400 ATAGTTCCTAAAGCGGAAAGCGTAAGTCTCTTATGGCACACCTCAGCCCAAGAAGATT 459
QY 162 PLeuCysHisMetAlaAsnGlyIleGluAspCysAspAspGln-----ProProLysG 180
Db 460 GATTATAGATAGAGAGACTGGGATCGT-GACAGGATTTTGAGCTCATGCTCTCTCTCT 513
QY 180 uLysLysSerLysSerAlaLysLysLysLysArgSerLysAlaLysGlnGluArgGluAl 200
Db 514 -----CCAGGAAGCAATCGGGAAGACCT 536
QY 200 aserProVal---GluPheAlaIleAspProAsnGluProThrTyrCysLeuCysAsnG 219
Db 537 TATGCCAATTGAAGACGACGACCAATCGATCCAAACGAACCACTTACTGTCTGCCATCA 596
QY 219 nValSerTyrGlyGluMetIleGlyCysAspAsnGluGlnCysPro---IleGluTrpPh 238
Db 597 GGTGCTCTTTGAGACATGATTCCTGTGACATGAGAAATTCCAAGGAGGTGAATGGTT 656
QY 238 eHisPheSerCysValSerLeuThrTyrLysPro-----LysGlyLysTrp 253
Db 657 TCACATATACATCGTTGGCTCACACCTGACACCATTCARAGGGAATGG 708

RESULT 12
US-09-801-368-257
; Sequence 257, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250Alman, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 257
; LENGTH: 993
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-257

Alignment Scores:
Pred. No.: 5,46e-21 Length: 993
Score: 263.50 Matches: 75
Percent Similarity: 37.04% Conservativity: 45

```



Best Local Similarity:	23.15%	Mismatches:	105
Query Match:	17.79%	Indels:	99
DB:	10	Gaps:	8

US-09-513-365A-1 (1-280), x US-09-801-368-257 (1-993)

QY	28	ValGlnAspTyrLeuGluCysValGlnSerLeuProHisAspMetGlnArgAsnValSer	47
Db	31	TTAAACGACATAACTCGAGTCGTGGAGAAATTCCACTGCCACCTCCAGATATTTAACA	90
QY	48	ValLeuArgGluLeuAspAsnLysTyrGlnGlnThrLeuLysGluIleaspaspValTyr	67
Db	91	TTACTACACGAATAAGATGCAAAATGTGTGCATTCTATGCCGAATTGTGAACGAGGAGA	150
QY	68	GluLysTyrLysLysGluAspAsp-----	LeuAsn 77
Db	151	GATAAGTTCTTCAAGAAAAAGACTTCAATAAAGATCACCAAACAAAGTAAGACTGCTCAAT	210
QY	78	GlnLysLysArgLeuGlnGlnLeuLeuGlnArgAlaLeulleAsnSerGlnGluLeuGly	97
Db	211	AATATCAACAAGATTATGAAGAATCATGTCATCGCTG-----	249
QY	98	AspGluLysIleGlnIleValThrGlnMetLeuLeuValGluAsnArgAlaArgGln	117
Db	250	GAGGAGAAATGCATGTCATCCATTATGCTGGATAATCTAGACAGATTGAGTCCCGG	309
QY	118	MetGluLeu-----	120
Db	310	TTGAATTTGGCGTGATGAAGTCGCAATCAAGAACACAGAAATTCCTAGAGGTTTTAAGACTG	369
QY	121	-----HisSerGlnCysPheGlnAspProAlaGluSerGluArgAlaSerAsp	136
Db	370	GGTGTGGACACCATCCAGCAATGCACCTCATCATGAATATGGAATAAATAAGAGAGC	429
QY	137	LysAlaLysMetAspSerSerGlnProGluArgSer---SerArgArgPro---	152
Db	430	AAATCAACAGCAAAATCGTCGACGGCACTGAAGAGCGCAATCAAGAGAGAACGCATGGCT	489
QY	153	-----ArgArgGln-----ArgThrSerGluSerArg	161
Db	490	GCCAACAGAGCAGCGGGCGGAACATTACTCCGCCAGCACACCAACAGCAGCTCAAG	549
QY	162	AspLeuCysHisMetalAsnGlyIleGluAspCysAspGlnProProLysGluLys	181
Db	550	AACGACGCAAACTCGGAGGCAGCAGGCAGCAGACCAACACACTGGTACACACACA	609
QY	182	LysSerLys-----SerAlaLysLys	188
Db	610	AACTCAAGAAAAAGACCGCAACGCTGCCAATACCAACAACGCCGATCCAGAGACCAAAA	669
QY	189	LysLysSerLysAlaLysGlnGluArgGluAlaSerProValGluPheAlaIle---	207
Db	670	CGCAAGAGAGAGTTGCCACCACACACCGCGTTTCACCAAGCACTATCAGCAGCGCAACTGCC	729
QY	207	-----	207
Db	730	GTCAATAATGGCAGGATAGGTATACATCTACACGCTCCAGGGAGTTAGCACCGTCGGAAC	789
QY	208	-----AspProAsnGluProThrTyrCys	215
Db	790	AGCAACAACAGCAGGATATCAAGACCAAAAAACCAACGACTACGCGCAACCGCTCTACTGC	849
QY	216	LeuCysAsnGlnValSerTyrGlyCluMetileglyCysAspAsnGluGlnCysProfile	235
Db	850	TACTGTAAACCAAGTGGCATACGGGAATGTGGGTGTGATGGCGCAGACTGTGAGCTA	909
QY	236	GluTrpPheHisPheSerCysValSerLeuThrTyrLysProLysGlyLysTryptyrCys	255
Db	910	GAATGGTTCCATTTCCTCATGTATTGGACTCGAAACTCTACTAAGGCCAAGTGGTATTGC	969
QY	256	ProLysCysArg	259
Db	970	GACGACTGCAAA	981

RESULT 13

```

US-09-918-995-16047
: Sequence 16047, Application US/09918995
: Publication No. US20030073623A1
:
: GENERAL INFORMATION:
:
:   APPLICANT: Hyseq, Inc.
:   TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
:   TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
:   FILE REFERENCE: 20411-756
:   CURRENT APPLICATION NUMBER: US/09/918,995
:   CURRENT FILING DATE: 2001-07-30
:   PRIOR APPLICATION NUMBER: US/09/235,076
:   PRIOR FILING DATE: 1999-01-20
:   NUMBER OF SEQ ID NOS: 38054
:   SOFTWARE: FastSEQ for Windows Version 3.0
:   SEQ ID NO 16047
:   LENGTH: 456
:   TYPE: DNA
:   ORGANISM: Homo sapiens
:
: FEATURE:
:   NAME/KEY: misc_feature
:   LOCATION: (1)...(456)
:   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-16047

```

Alignment Scores:		
Pred. No.:	1.72e-09	Length:
Score:	160.00	Matches:
Percent Similarity:	54.69%	Conservative:
Best Local Similarity:	30.47%	Mismatches:
Query Match:	10.80%	Indels:
DB:	9	Gaps:
		3
		456

US-09-513-365A-1 (1-280) x US-09-918-995-16047 (1-456)

Qy	27	TyrValGlnAspTyrLeuGluCysValGluSerLeuProHisAspMetGlnArgAsnVal	46
Db	59	TACTTTGGACACTATCTCGACAGTATCGAGAACCTTCCTCGCGAACTTCAGAGGAACCTTC	118
Qy	47	SerValLeuArgGluLeuAspAsnLysTyrGlnGluThrLeuLysGluIleAspAspVal	66
Db	119	CAGCTGATCGAGAGCTCGACCAGAGCGGGAAGATAGAAACGACAGATTGACATCCGT	178
Qy	67	TyrGluLysTyr---LysLysGluAspAspLeuAsnGlnLysLysArgLeuGlnGlnLeu	85
Db	179	GCTCGACAGTACATCTCCACGGTGAAGACGCTGTCTCCAGACACAGCGCGTGGAGCGCCTG	238
Qy	86	-----LeuGlnArgAlaLeuIleAsnSerGlnGluLeuGlyAspGluLysIleGlnIle	103
Db	239	CAGAAGATCCAGACGCCCTACAGCAAGTGCAGAGGAATACAGTGACGACAAAAGTGCAGCTG	298
Qy	104	ValThrGlnMetLeuGluLeuValGluAsnArgAlaArgGlnMetGluLeuHisSerGln	123
Db	299	GCCATGCAGACCTACGAGATGGTGATTAACATCTCGAAGCGCTTGATCGACACCTGGCGG	358
Qy	124	CysPheGlnAspProAlaGluSerGluArgAlaSerAspLysAlaLysMetAspSerSer	143
Db	359	CGCTTTTGAA-----GCAGATCTCAAGGACAAAGATGGAGGCGAGT	397
Qy	144	GlnProGluArgSerSerArgArg	151
Db	398	GATTTTGAAGAGCTCCGGAGGCGCA	421

## RESULT 14

```

US-10-044-090-370
; Sequence 370, Application US/10044090
; Patent No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090

```

; CURRENT FILING DATE: 2002-01-09  
; NUMBER OF SEQ ID NOS: 850  
; SOFTWARE: PERL Program  
; SEQ ID NO 370  
; LENGTH: 7805  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: Incyte ID No. US20020137081A1 235885.5  
US-10-044-090-370

Alignment Scores:  
Pred. No.: 0.000267 Length: 7805  
Score: 130.50 Matches: 46  
Percent Similarity: 38.60% Conservative: 20  
Best Local Similarity: 26.90% Mismatches: 66  
Query Match: 8.81% Indels: 39  
DB: 12 Gaps: 8

US-09-513-365A-1 (1-280) x US-10-044-090-370 (1-7805)

```
QY 122 SerGlnCysPheGlnAspProAlaGluSerGluArgAlaSerAspLysAlaLysMetAsp 141
DB 1416 TCGTATGTTTTCAGAGCGACGAAGTCTGAACCCAGAGGCTGAGGAATCAGACCTGGAC 1475
QY 142 SerSerGlnProGluArgSerArgArgProArgArgGlnArgThrSerGluSerArg 161
DB 1476 AGTGGCAGTGTCCACAGTCCCTCAGCCGCGCCT- 1508
QY 162 AspLeuCysHisMetAlaAsnGlyLeuGluAspCysAspAspGlnProProLysGluLys 181
DB 1509 -----GATGGCCCTGTCCGCCACCAAG 1529
QY 182 LysSerLysSerAlaLysLysLysLysArgSerLysAlaLysGlnGluArgGluAlaSer 201
DB 1530 AAACCTAAAGAGAGCGCGCCAGGAAGGAAGAGAGAGTCTGTGGCTGTCTGCAGTG 1589
QY 202 ProValGluPheAlaIleAspProAsnGlu-----ProThrTyrCys---LeuCys 217
DB 1590 GCGGGGAGGAGGAGGTGATGGCTACGAGCGGATCACCAGATTACTGTGAGGTGTC 1649
QY 218 AsnGlnValSerTyrGlyGluMetIleGlyCysAspAsnGluCysProIleGluTrp 237
DB 1650 CAGCAG-----GGTGGGAAATATTCTGTGTGACACC-----TGCCCT---CGTGCC 1694
QY 238 PheHisPheSerCysVal-----SerLeuThrTyrLysProLysGlyLysTrpTyrCys 255
DB 1695 TACCACCTCTGTCTGCTTGCCTGATCCTGAGCTTGACCGGGCTCCAGAGGGCAAAATGGAGCTGC 1754
QY 256 ProLysCysArgGlyAsp-----AsnGluLysThrMetAspLysSer 269
DB 1755 CCTCACTGTGAGAGGAGGGGTCCAGTGGGAGGCCAAGAGGAAGAAGAAATCGAAG 1814
QY 270 ThrGluLysThrLysLysAspArgArgSerArg 280
DB 1815 AGGAGGAGGAGGAAGGAGGAGGAGGAGGAGG 1847
```

RESULT 15

US-09-764-864-471  
; Sequence 471, Application US/09764864  
; Patent No. US20020132753A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT223  
; CURRENT APPLICATION NUMBER: US/09/764,864  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1792  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 471  
; LENGTH: 915

; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (894)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (903)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (905)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (914)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-09-764-864-471

Alignment Scores:  
Pred. No.: 1.6e-05 Length: 915  
Score: 129.50 Matches: 34  
Percent Similarity: 40.77% Conservative: 19  
Best Local Similarity: 26.15% Mismatches: 34  
Query Match: 8.74% Indels: 43  
DB: 10 Gaps: 7

US-09-513-365A-1 (1-280) x US-09-764-864-471 (1-915)

```
QY 166 MetAlaAsnGlyLeuGluAspCys-----Asp 174
DB 190 CTGGCGGGGTCTGTGAGGAGTGCCTGGCGGAGCGGAGATATGAGGGAGATGTTTCAGAC 249
QY 175 AspGlnProPro-----LysGluLysLysSerLysSerAlaLysLysLysArg 191
DB 250 CCAGAGCTCCAGATGTCGGGGAGGACAGCAAGTCCGAGAATGGGAG-----297
QY 192 SerLysAlaLysGlnGluArgGluAlaSerProValGluPheAlaIleAspProAsnGlu 211
DB 298 -----AATGCG 303
QY 212 ProThrTyrCysLeuCysAsnGlnValSerTyrGly---GluMetIleGlyCysAspAsn 230
DB 304 CCCATCTACTGCATCTGCGCAACCGGACATCAACTGCTTCATGATCGGGTGTGACAAC 363
QY 231 GluGlnCysProIleGluTrpPheHisPheSerCysValSerLeuThrTyrLysProLys 250
DB 364 -----TGCAAT---GAGTGGTTCATGGGGACTGCATCCGGATCCTGAGAAGATGGCC 414
QY 251 -----GlyLysTrpTyrCysProLysCysArgGlyAspAsnGluLysThrMetAsp 267
DB 415 AAGGCCATCCGGGAGTGTGTCTGGGAGTGCAGAGAGAAAGACCCCAAGCTAGAGATT 474
QY 268 LysSerThrGluLysThrLysLysAspArg 277
DB 475 CGCTATCGGCACAAAGAATGTCACGGGAGCGG 504
```

Search completed: June 20, 2003, 04:42:50  
Job time : 173 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003. Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2003, 01:21:09 ; Search time 80 Seconds  
(without alignments)  
4140.136 Million cell updates/sec

Title: US-09-513-365A-2  
Perfect score: 1080  
Sequence: 1 gcggccgcgcgcgtgcatg.....aaaaaaaaaaaaaaaaaa 1080

Scoring table: IDENTITY\_NUC  
Gapop 10,0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTCUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1046	96.9	1078	US-09-601-478-7	Sequence 7, Appli
2	838.4	77.6	840	US-09-601-478-6	Sequence 6, Appli
3	825.2	76.4	1154	US-09-484-970B-81	Sequence 81, Appli
4	237.4	22.0	873	US-09-006-783A-4	Sequence 4, Appli
5	183.8	17.0	1902	US-09-258-257-1	Sequence 1, Appli
6	183.8	17.0	1902	US-09-258-371-1	Sequence 1, Appli
7	183.8	17.0	1902	US-09-258-721A-1	Sequence 1, Appli
8	183.8	17.0	1902	US-08-751-230-1	Sequence 1, Appli
9	183.8	17.0	1902	US-09-499-082-1	Sequence 1, Appli
10	183.8	17.0	1902	US-09-258-372-1	Sequence 1, Appli
11	183.8	17.0	2061	US-09-258-371-9	Sequence 9, Appli
12	183.8	17.0	2061	US-08-751-230-9	Sequence 9, Appli
13	183.8	17.0	2061	US-09-499-082-9	Sequence 9, Appli
14	183.8	17.0	2061	US-09-258-372-9	Sequence 9, Appli
15	182.2	16.9	2061	US-09-006-783A-2	Sequence 2, Appli
16	182.2	16.9	2061	US-09-159-871-1	Sequence 1, Appli
17	174.4	16.1	633	US-09-006-783A-6	Sequence 6, Appli
18	95.6	8.9	678	US-09-195-286-3	Sequence 3, Appli
19	95.6	8.9	699	US-09-195-286-2	Sequence 2, Appli
20	83.6	7.7	451	US-09-370-838-146	Sequence 146, App
21	72.4	6.7	7218	US-08-232-463-14	Sequence 14, Appli
22	50.6	4.7	860	US-07-847-010-18	Sequence 18, Appli
23	49	4.5	593	US-09-385-982-262	Sequence 262, App
24	46.8	4.3	6124	US-08-213-419B-3	Sequence 3, Appli
25	46.2	4.3	1189	US-07-781-034-4	Sequence 4, Appli
26	46.2	4.3	1189	PCT-US92-08328-4	Sequence 4, Appli
27	45.8	4.2	19124	US-08-487-826B-13	Sequence 13, Appli

28	43.8	4.1	3489	2	US-08-728-323A-1	Sequence 1, Appli
29	43.8	4.1	3489	4	US-09-298-568-1	Sequence 1, Appli
c 30	43.8	4.1	32207	2	US-08-770-379-20	Sequence 20, Appli
c 31	43.8	4.1	32207	4	US-08-757-659A-20	Sequence 20, Appli
c 32	43.8	4.1	32207	4	US-09-230-371A-20	Sequence 20, Appli
33	43.4	4.0	1147	1	US-08-665-716-1	Sequence 1, Appli
34	43.2	4.0	3825	4	US-09-208-742-3	Sequence 3, Appli
35	43	4.0	3483	4	US-09-130-491-3	Sequence 3, Appli
c 36	42.8	4.0	603	4	US-09-385-982-251	Sequence 251, App
c 37	42.8	4.0	774	3	US-08-956-307B-11	Sequence 12, Appli
38	42.8	4.0	778	3	US-08-956-307B-11	Sequence 11, Appli
39	42.8	4.0	3377	6	5198345-16	Patent No. 5198345
40	42.2	3.9	841	4	US-09-004-731-40	Sequence 40, Appli
c 41	42.2	3.9	841	4	US-09-004-731-42	Sequence 42, Appli
42	42.2	3.9	841	4	US-09-032-215-46	Sequence 46, Appli
c 43	42.2	3.9	841	4	US-09-032-215-48	Sequence 48, Appli
44	42.2	3.9	841	4	US-08-749-699-40	Sequence 40, Appli
c 45	42.2	3.9	841	4	US-08-749-699-42	Sequence 42, Appli

## ALIGNMENTS

RESULT 1  
US-09-601-478-7  
; Sequence 7, Application US/09601478  
; Patent No. 6403785  
; GENERAL INFORMATION:  
; APPLICANT: Otsuka Pharmaceutical Co., Ltd.  
; TITLE OF INVENTION: Human TSC403 gene and human INGL gene  
; FILE REFERENCE: Q60193  
; CURRENT APPLICATION NUMBER: US/09/601,478  
; CURRENT FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: JP H10-134679  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: JP H10-73234  
; PRIOR FILING DATE: 1998-03-05  
; PRIOR APPLICATION NUMBER: JP H10-38133  
; PRIOR FILING DATE: 1998-02-03  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 1078  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: human embryonic brain cDNA library  
; NAME/KEY: CDS  
; LOCATION: (92)..(931)  
US-09-601-478-7

Query Match	96.9%	Score 1046;	DB 4;	Length 1078;
Best Local Similarity	99.5%	Pred. No. 1.4e-264;		
Matches 1049;	Conservative	0;	Mismatches 5;	Indels 0;
Gaps	0;			
Qy	1	CGCGCGCGCGCGGTGTCATGTCCGCTGCTCGATCGGAGCGCGGCGCGCGGAT	60	
Db	25	CGCGCGCGCGCGGTGTCATGTCCGCTGCTCGATCGGAGCGCGGCGCGCGGAT	84	
Qy	61	CGGCGAGGATTTAGGGCAGCAGCAGCAGCAGTGTCTGCTCGGCTGGCTCTGCTGACCGG	120	
Db	85	CGGCGAGGATTTAGGGCAGCAGCAGCAGCAGTGTCTGCTCGGCGCGGCTCTGACCGG	144	
Qy	121	GGAGCGGAGCGGCTGCTCCTGCTAGTGCAGGACTACCTTGTAGTGGTGGAGTCGCT	180	
Db	145	GGAGCGGAGCGGCTGCTCCTGCTAGTGCAGGACTACCTTGTAGTGGTGGAGTCGCT	204	
Qy	181	GCCCGACACATGAGAGGAACTGCTGCTGCTGAGAGCTGGACAAATATCAAGA	240	
Db	205	GCCCGACACATGAGAGGAACTGCTGCTGCTGAGAGCTGGACAAATATCAAGA	264	
Qy	241	AACGTTAAAGAAATTTGATGTCTACGAAATATTAAGAAAGAGATGATTTAAACCA	300	

Db	265	AACGTTAAAGGAATTTGATGATCTCTACGAAAAATATATAGAAAGAGATGATTTAAACCA	324
Qy	301	GAAGAAACGCTTACAGCAGCTTCTCCAGAGAGACACTAAATTAATAGTCAAGAAATGGGAGA	360
Db	325	GAAGAAACGCTTACAGCAGCTTCTCCAGAGAGACACTAAATTAATAGTCAAGAAATGGGAGA	384
Qy	361	TGAAAAATACAGATTTGTTACACAAATGCTCGAATTTGGTGGAAAAATTCGGGCAAGACAAAT	420
Db	385	TGAAAAATACAGATTTGTTACACAAATGCTCGAATTTGGTGGAAAAATTCGGGCAAGACAAAT	444
Qy	421	GGAGTTACACTCACAGTGTTTTCCAAGATCTCTGTAAGTGAACGAGGCCTTCAGATAAAGC	480
Db	445	GGAGTTACACTCACAGTGTTTTCCAAGATCTCTGTAAGTGAACGAGGCCTTCAGATAAAGC	504
Qy	481	AAAGATGGATTCCAGCCCAACGAGAAAGATCTTCAAGAAGACCCCGCAGGCGAGCGACCAG	540
Db	505	AAAGATGGATTCCAGCCCAACGAGAAAGATCTTCAAGAAGACCCCGCAGGCGAGCGACCAG	564
Qy	541	TGAAGCCGTGATTTATGTCACATGGCAATGGGATTGAAGACTGTGATGATCAGCCACC	600
Db	565	TGAAGCCGTGATTTATGTCACATGGCAATGGGATTGAAGACTGTGATGATCAGCCACC	624
Qy	601	TAAAGAAAAGAAATCCAAGTCACCAAGAAAAAGAAACGCTCCAAGGCCCAAGCAGGAAAG	660
Db	625	TAAAGAAAAGAAATCCAAGTCACCAAGAAAAAGAAACGCTCCAAGGCCCAAGCAGGAAAG	684
Qy	661	GGAAGCTTCACCTGTTGAGTTTGCATCTAGATCCTTAATGAACCTACACTGCTTATGCAA	720
Db	685	GGAAGCTTCACCTGTTGAGTTTGCATCTAGATCCTTAATGAACCTACACTGCTTATGCAA	744
Qy	721	CCAAGTGTCTTATGGGGAGATGATAGGATGTGACATGAACAGTGTCCAATTGCAATGGTT	780
Db	745	CCAAGTGTCTTATGGGGAGATGATAGGATGTGACATGAACAGTGTCCAATTGCAATGGTT	804
Qy	781	TCACTTTTCATGTGTTTCACTTACCTATAAACCAAGGGGAAATGGTATTGCCCAAGTG	840
Db	805	TCACTTTTCATGTGTTTCACTTACCTATAAACCAAGGGGAAATGGTATTGCCCAAGTG	864
Qy	841	CAGGGAGATAATGAGAAAAACAATGGACAAAAGTACTGAAAGACAAAAAGATAGAAG	900
Db	865	CAGGGAGATAATGAGAAAAACAATGGACAAAAGTACTGAAAGACAAAAAGATAGAAG	924
Qy	901	ATCGAGGTAGTAAGGCCATCCACATTTTAAAGGGTATTGTCTTTTATATATATTCGTT	960
Db	925	ATCGAGGTAGTAAGGCCATCCACATTTTAAAGGGTATTGTCTTTTATATATATTCGTT	984
Qy	961	TGCTTTTCAGAAAAATGTTTAAAGGTAAATGCATGAAGACTATGCAATATTTTTTATCATTA	1020
Db	985	TGCTTTTCAGAAAAATGTTTAAAGGTAAATGCATGAAGACTATGCAATATTTTTTATCATTA	1044
Qy	1021	GTATTAAATGGTGATTAAAAAGTTTGTGTACTTTG	1054
Db	1045	GTATTAAATGGTGATTAAAAAGTTTGTGTACTTTG	1078

## RESULT 2

RESOLUT 2  
US-09-601-478-6  
; Sequence 6, Application US/09601478  
; Patent No. 6403785  
; GENERAL INFORMATION:  
; APPLICANT: Ostuka Pharmaceutical Co., Ltd.  
; TITLE OF INVENTION: Human TSC403 gene and human INGIL gene  
; FILE REFERENCE: Q60193  
; CURRENT APPLICATION NUMBER: US/09/601,478  
; CURRENT FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: JP H10-134679  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: JP H10-73234  
; PRIOR FILING DATE: 1998-03-05  
; PRIOR APPLICATION NUMBER: JP H10-38133  
; PRIOR FILING DATE: 1998-02-03  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.1



168	DB	CTACGAGCGCTTCAGTCCGGAGACAGAGCGGGCGCAGAGCGCGGATGCTGCACGTGT	227
325	QY	CCAGAGAGCACTAATTAATAGTCAAGATTTGGGAGATGAAAAAATACAGATTTGTTACACA	384
228	DB	GCAGCGCGGCTGATCCCGACCGAGAGCTGGGCGAGAGAAATCCAGATCGTGAGCCA	287
385	QY	AATGCTGAATGGTGGGAAATCGGGCAAGACAAATGGAGTTACACTACAGTGTTCOA	444
288	DB	GATGTTGGAGCTGGTGGAGAACCGCACGCGCAGGTGGACAGCCACGTGGAGCTGTTCGA	347
445	QY	AGATC-----CTGCTGAAAGTGAACGAGCCTCAGATA	476
348	DB	GGCGCAGCAGGAGCTGGCGGACACAGTGGCAACAGCGCGCAAGTTGGCGCGCACAGGCC	407
477	QY	AAGCAAAAGATGGATTCAGGCCAACACAGAAAGATCTTCAAGAACACACCCCGCAG	534
408	DB	CAATGGCGATCGGTAGCCGAGTGTGACAGCCCAACAGCAACGGCTACAGGGCGCAGCG	467
535	QY	GACCACTGAAAGCCGTGATTTATGTCATGTGGCAATGGGATTTGAAGACTGTGATGATCA	594
468	DB	CAACAACGAGAACCGTGAGAACCGCTCAGCAACACGACCACGACGCGCGCCTCGGG	527
595	QY	GCCACCTTAAAGAAAGAAATCCAGTCCAGCAAGAAAGAAACGGTCCAGGCCCAAGCA	654
528	DB	CACACCCAAGGAGAAGAGGCCAAGACTCTCAAGAAGAAGACGGCTCCAAGGCCCAAGGC	587
655	QY	GGAAAGGGAAGCTTCACCTGTTGAGTTTGAATAGATCCTTAATGAACCTACATACGTCT	714
588	DB	GGAGCGAGCGGCTCCCTGCCGACCTCCCTCATCGACCCCAACCAACCACTACTGTCT	647
715	QY	ATGCAACCAAGTGTCTTATGGGGAGATGATAGGATGTGCAATGAACAGTGTCCCAATGA	774
648	DB	GTGCAACCAAGTCTCTTATGGGAGATGATCGGCTGCGCAACGACGAGTGCCCATCGA	707
775	QY	ATGGTTTTCACATTTCATGTGTTTTCACATTACCTATAAACCAAGGGGAATGTGTTGCC	834
708	DB	GTGGTTCACATTCTGTCGCTGGGGCTCAATCATAAACCCCAAGGGCAAGTGGTACTGTCC	767
835	QY	AAAGTGCAGGGAGATTAATGAGAAACAAATGGCAAAAAGTACTGAAAACACAAAAAAGA	894
768	DB	CAAGTCCCGGGGAGACGAGAGACCATGGACAAAGCCCTGGAGAAATCCAAAAAAGA	827
895	QY	TAG 897	
828	DB	GAG 830	

## RESULT 5

```

US-09-258-257-1
Sequence 1, Application US/09250257
Patent No. 5965398
GENERAL INFORMATION:
APPLICANT: GARKAVTSEV, Igor
APPLICANT: RIABOMOL, Karl
TITLE OF INVENTION: DNA SEQUENCE ENCODING A TUMOR
SUPPRESSOR GENE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESS: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/258,257
FILING DATE:

```



LOCATION: 109..738  
US-08-569-721A-1

Query Match 17.0%; Score 183.8; DB 3; Length 1902;

Best Local Similarity 57.0%; Pred. No. 1.4e-38;  
Matches 394; Conservative 0; Mismatches 267; Indels 30; Gaps 2;

```
QY 237 AAGAAACGTTAAGAGAAATGATGCTACGAAATAATAGAAAGAGATGATTAA 296
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 35 AACAGATCTGAAGAGCTAGAGAGTCTACGAGCGCTTCAGTCGCGAGACAGCGGG 94

QY 297 ACAGAAAGAACTCTACAGAGCTTCTCCAGAGACACTAATTAATAGTCAAGAA 356
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 95 CGCAGAGGGGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 154

QY 357 GAGATGAAATAATACAGATGTTTACACAAATGCTCAATTTGTTGGAATCGGCA 416
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 155 GCGACAGAGATCCAGATCTGAGCCAGATGTTGGAGCTGTTGGAGAACCCGCGC 214

QY 417 AATGAGTTTACACTCACAGTGTTCCTCAAGTC----- 449
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 215 AGGTGACAGCCAGCTGAGCTGTTGAGCGCGACGAGAGCTGGCGGACAGTGGGCA 274

QY 450 -CTGTGAAGTGAACGAGCTCTACAGAGCTTCTCCAGAGACACTAATTAATAGTCA 508
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 275 ACAGCGGAGAGTTGGCGGCGAGAGCGGCAATGGCGATCGGTAGCGAGCTGACAGC 334

QY 509 TCTTCAAGAA--GACCCCGCAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 566
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 335 CCACAGCAAGCGCTCACGCGCGAGCGGCAACAGCAACCGTGAGAACCGTCTGACA 394

QY 509 TCTTCAAGAA--GACCCCGCAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 566
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 335 CCACAGCAAGCGCTCACGCGCGAGCGGCAACAGCAACCGTGAGAACCGTCTGACA 394

QY 567 CAATGGGATTAAGACTGTGATGATCAGCCACCTTAAGAAAGAAATCAAGTCAAG 626
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 395 ACCAGCACACGAGCGCGCTCGGGCACACCCCAAGGAGAGCGGCAAGACCTCCA 454

QY 627 TAGATCTAATGAACCTACATCTGTTATGCAACCAAGTCTTATGGGAGATGATAG 746
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 515 TCGACCCCAAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 574

QY 747 GATGTGACAAATGAACAGTGTCCAAATGATGTTTCACTTTTCACTTTTCACTTAC 806
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 575 GCTGCGACAAACGAGTGTCCCATCGAGTGTGCTTCTGCTGCTGCTGCTGCTGCT 634

QY 807 ATAAACCAAGGGGAATGTTGCTTCCCAAGTGTGAGGAGATATGAGAAACAAATGG 866
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 635 ATAAACCAAGGGGAATGTTGCTTCCCAAGTGTGAGGAGATATGAGAAACAAATGG 866

QY 867 ACAAAAGTACTGAAAGACAAAAAGGATAG 897
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 695 ACAAGCCCTGGAGAAATCCAAAAAGAGAG 725
```

## RESULT 8

US-08-751-230-1

Sequence 1, Application US/08751230

Patent No. 6117633

GENERAL INFORMATION:

APPLICANT: Garkavtsev, Igor

APPLICANT: Riabowol, Karl

TITLE OF INVENTION: DNA SEQUENCE ENCODING THE TUMOR

TITLE OF INVENTION: SUPPRESSOR GENE IN1

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: 699 Prince Street

CITY: Alexandria

STATE: VA

COUNTRY: USA

ZIP: 22313-1404

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/751,230  
FILING DATE: 15-NOV-1996  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/569721  
FILING DATE: 08-DEC-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Mooi, Leslie A.  
REGISTRATION NUMBER: 37,047  
REFERENCE/DOCKET NUMBER: 028722-144  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-854-7400  
TELEFAX: 415-854-8275  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1902 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cdna  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 109..741  
US-08-751-230-1

```
Query Match 17.0%; Score 183.8; DB 3; Length 1902;
Best Local Similarity 57.0%; Pred. No. 1.4e-38;
Matches 394; Conservative 0; Mismatches 267; Indels 30; Gaps 2;

QY 237 AAGAAACGTTAAGAGAAATGATGCTACGAAATAATAGAAAGAGATGATTAA 296
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 35 AACAGATCTGAAGAGCTAGAGAGTCTACGAGCGCTTCAGTCGCGAGACAGCGGG 94

QY 297 ACAGAAAGAACTCTACAGAGCTTCTCCAGAGACACTAATTAATAGTCAAGAA 356
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 95 CGCAGAGGGGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 154

QY 357 GAGATGAAATAATACAGATGTTTACACAAATGCTCAATTTGTTGGAATCGGCA 416
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 155 GCGACAGAGATCCAGATCTGAGCCAGATGTTGGAGCTGTTGGAGAACCCGCGC 214

QY 417 AATGAGTTTACACTCACAGTGTTCCTCAAGTC----- 449
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 215 AGGTGACAGCCAGCTGAGCTGTTGAGCGCGACGAGAGCTGGCGGACAGTGGGCA 274

QY 450 -CTGTGAAGTGAACGAGCTCTACAGAGCTTCTCCAGAGACACTAATTAATAGTCA 508
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 275 ACAGCGGAGAGTTGGCGGCGAGAGCGGCAATGGCGATCGGTAGCGAGCTGACAGC 334

QY 509 TCTTCAAGAA--GACCCCGCAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 566
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 335 CCACAGCAAGCGCTCACGCGCGAGCGGCAACAGCAACCGTGAGAACCGTCTGACA 394

QY 567 CAATGGGATTAAGACTGTGATGATCAGCCACCTTAAGAAAGAAATCAAGTCAAG 626
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 395 ACCAGCACACGAGCGCGCTCGGGCACACCCCAAGGAGAGCGGCAAGACCTCCA 454

QY 627 TAGATCTAATGAACCTACATCTGTTATGCAACCAAGTCTTATGGGAGATGATAG 686
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 455 TCGACCCCAAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 514

QY 687 TAGATCTAATGAACCTACATCTGTTATGCAACCAAGTCTTATGGGAGATGATAG 746
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 515 TCGACCCCAAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 574

QY 747 GATGTGACAAATGAACAGTGTCCAAATGATGTTTCACTTTTCACTTTTCACTTAC 806
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 575 GCTGCGACAAACGAGTGTCCCATCGAGTGTGCTTCTGCTGCTGCTGCTGCTGCT 574

QY 747 GATGTGACAAATGAACAGTGTCCAAATGATGTTTCACTTTTCACTTTTCACTTAC 806
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```





```

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mooi, Leslie A.
; REGISTRATION NUMBER: 37,047
; REFERENCE/DOCKET NUMBER: 028722-144
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-854-7400
; TELEFAX: 415-854-8275
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1902 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 109..741
; US-09-258-372-1

Query Match 17.0%; Score 183.8; DB 4; Length 1902;
Best Local Similarity 57.0%; Pred. No. 1.4e-38;
Matches 394; Conservative 0; Mismatches 267; Indels 30; Gaps 2;

QY 237 AGAAGAGGTTAAAGGAAATGATGTTCTACGAAATATATAAGAAAGATGATTAA 296
Db 35 AACAGATCTGAAGAGCTAGACGAGTGTCTACGAGCGCTTCAGTCGCGAGACAGACGGG 94
QY 297 ACCAGAGAAACCTCTACGAGCTTCTCCAGAGACACTAATATAGTCAAGATTGG 356
Db 95 CGCAGAGCGCGGATGCTGCACTGTGCGAGCGCGCTGTATCCGCGCAGGAGCTGG 154
QY 357 GAGATGAAATAATACAGATGTTTACACAAATGCTCGAATTTGGTGGAAATCGGGCAAGAC 416
Db 155 GCGACGAGAGATCCAGATCGTGTGAGCCAGTGTGTGAGCTGTGTGGAGACCGACCGGC 214
QY 417 AATGGAGTTTACTCAGCTCAGCTGTTTCCAGATC----- 449
Db 215 AGGTGACGACCGACGCTGGAGCTTTCGAGCGCGAGGAGCTGGCGCAGACAGTGGGCA 274
QY 450 -CTGCTGAAGTGAACGAGCTCAGATAAGCAAGATGATTCACGCGCAAGCAAGA 508
Db 275 ACAGCGCGAAGGTTGGCGGAGCAGCGCCCAATGGCGATGGGTAGCGCAGCTGTGACAGC 334
QY 509 TCTTCAAGAA--GACCCCGCAGCGCAGCGACAGTGAAGCCGCTGATTTATGTCACATGG 566
Db 335 CCAACAGCAAGCGCTCAGCGCGCAGCGCAACCAACGAGAACCGTGAGAACCGCTCCAGCA 394
QY 567 CAATGGGATTGAGACTGTGATGATCAGCCACCTAAGAAAGAAATCCAGTACAGCA 626
Db 395 ACCACGACCGACGAGCGCGCTCGGCGCACACCCAGGAGAGAAAGGCGCAAGACCTCCA 454
QY 627 AGAAAGAAACGCTCCAGGCGCAAGCAAGGAGGAGCTTCACCTGTTGAGTTTGCA 686
Db 455 AGAAGAGAGCGCTCCAGGCGCAAGCGGAGCGGAGCGGCTCCCTGCGGACCTCCCA 514
QY 687 TAGATCCTAATGAACCTACATCTGTTATGCAACCAAGTGTCTTATGGGAGATGATAG 746
Db 515 TCGACCCCAACGAACCACTGCTGTGTGCAACCAAGTGTCTCTATGGGAGATGATCG 574
QY 747 GATGTGACATGACAGTGTCCAAATGAATGTTTTCATCTTTTCATCTTTTCACTTACCT 806
Db 575 GCTGCGACACGACGAGTGTCCCATGAGTGTGTTCACCTTCTGTCGCTGGGCTCAATC 634
QY 807 ATAAACCAAGGGGAATGTTGTCGCAAGTGCAGGGGAGATATAGAAACAACTGG 866
Db 635 ATAAACCAAGGGGAATGTTGTCGCAAGTGCAGGGGAGATATAGAAACAACTGG 866
QY 867 ACAAAGTACTGAAAGACAAAAAGGATAG 897
Db 695 ACAAAGCTTGGAGAAATCCAAAAAGAGAG 725

```

```

RESULT 11
US-09-258-371-9
; Sequence 9, Application US/09258371
; Patent No. 5986078
; GENERAL INFORMATION:
; APPLICANT: Garkavtsev, Igor
; APPLICANT: Riabowol, Karl
; TITLE OF INVENTION: DNA SEQUENCE ENCODING THE TUMOR
; TITLE OF INVENTION: SUPPRESSOR GENE IN1
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: 699 Prince Street
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-1404
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/258,371
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/751,230
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mooi, Leslie A.
; REGISTRATION NUMBER: 37,047
; REFERENCE/DOCKET NUMBER: 028722-144
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-854-7400
; TELEFAX: 415-854-8275
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2061 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16..900
; US-09-258-371-9

```

```

Query Match 17.0%; Score 183.8; DB 2; Length 2061;
Best Local Similarity 57.0%; Pred. No. 1.4e-38;
Matches 394; Conservative 0; Mismatches 267; Indels 30; Gaps 2;

QY 237 AAGAAACGTTAAAGGAAATGATGTTCTACGAAATATATAAGAAAGATGATTAA 296
Db 194 ACAGATCTGAGGAGCTAGACGAGTGTCTACGAGCGCTTCAGTCGCGAGACAGACGGG 253
QY 297 ACCAGAGAAACGCTCTACAGCAGCTTCTCCAGAGAGCACTAATTAATAGTCAAGAAATGG 356
Db 254 CGCAGAGCGCGGATGCTGCACTGTGTGCGAGCGCGCTGATCCCGCAGGAGCTGG 313
QY 357 GAGATGAAATAATACAGATGTTTACACAAATGCTCGAATTTGGTGGAAATCGGGCAAGAC 416
Db 314 GCGACGAGAAGATCCAGATCGTGTGAGCCAGTGTGTGAGCTGTGTGGAGAACCGCACGGC 373
QY 417 AATGGAGTTTACTCAGCTCAGCTGTTTCCAGATC----- 449
Db 374 AGGTGAGACAGCGCGCTGGAGCTTTCGAGCGCGAGGAGCTGGCGCAGACAGTGGGCA 433
QY 450 -CTGCTGAAGTGAACGAGCTCAGATAAGCAAGATGATTCACGCGCAACCAAGCAAGA 508
Db 434 ACAGCGCGAAGGTTGGCGGAGCAGCGCCCAATGGCGATCGGTAGCGCAGTCTGACAGC 493
QY 509 TCTTCAAGAA--GACCCCGCAGCGGAGGAGTGTGATTTATGTCACATGG 566

```

```

; LOCATION: 16..900
US-08-751-230-9

Query Match      17.0%; Score 183.8; DB 3; Length 2061;
Best Local Similarity 57.0%; Pred. No. 1.4e-38;
Matches 394; Conservative 0; Mismatches 267; Indels 30; Gaps 2;

QY 237 AAGAAACGTTAAAGGAAATTTGATGATCTCTACGAAAAATATAAGAAAGAAAGATGATTAA 296
Db 194 AACAGATCTCTGAAGGAGCTAGACAGTGTCTACGAGCGCTTCAGTCGCGAGACAGACGGG 253
QY 297 ACAGAGAAGAAACGCTCTACACAGCTTCTCAGACAGCACATTAATTAGTCAAGANTTGG 356
Db 254 CGCAGAAGCGCGGATGCTGCACCTGTGTGACGCGCGCTGATCCGACGCCAGGAGCTGG. 313
QY 357 GAGATGAAAAAATACAGATTGTTACACAAATGCTCGAATTGGTGGAAAAATCGGCAAGAC 416
Db 314 GCACGAGAAGATCCAGATCGTGAGCAGATGGTGAGCTGGTGAGAACCCGACGCGGC 373
QY 417 AAATGGAGTTTACACTCACAGTGTGTTCCAAAGATC----- 449
Db 374 AGTGGACACCCAGCTGGAGCTGTTCCAGGCGCAGCAGGAGCTGGCGCACAGTGGCA 433
QY 450 -CTGCTGAAAGTGAACGAGCGCTCAGATAAAGCAAAAGATGGATCCAGCCAACAGAAAGA 508
Db 434 ACAGCGCAAGGTTGGCGGACAGGCGCCAAATGGCGATGGCGTAGCGCAGTCTGACAAGC 493
QY 509 TCTTCAAGAA - GACCCCGCAGCAGCGGACCAGTGAAGACCGTGATTTATGTCACATGG 566
Db 494 CCAACAGCAAGCGCTCACGGCGGCGACGCAACACAGAGACCGGTGAGAACCGGTCACGA 553
QY 567 CAAATGGGATTAAGACTGTGATGATCAGCCACCTAAAGAAAAAATCCTCAAGTCAGCAA 626
Db 554 ACCACGACCAGCAGCGCGCGCTCGGGCACACCCCAAGGACGAGGCCAAGACCTCCA 613
QY 627 AGAAAAAGAAACGCTCCAAGGCCAAGCAGGAAGGGAAGCTTCACCTGTTGAGTTTGCAG 686
Db 614 AGAAGAAGAGCGCTCCAAGSCCAAGCGGAGCGAGGCGTCCCTGCCGACCTCCCA 673
QY 687 TAGATCCTAATGAACCTACATCTGCTTTATGCAACCAAGTGCTTTATGGGGAGATGATAG 746
Db 674 TCGACCCCAACGAAACCCACGTACTGTGTGTGCAACACCAAGGTCTCTATGGGGAGATGATCG 733
QY 747 GATGTGCAATGAACAGTGTCCAAATCAATGGTTTCACTTTTCATGTGTTTCACTTACCT 806
Db 734 GCTGGACACGACGAGTGGCCCATCGAGTGGTTCCACTTCTCTGGGTGGGGCTCAATC 793
QY 807 ATAAACCAAGGGGAAATGTTATTCGCCAAAGTGCAGGGGAGATAAATGAAAAACAATGG 866
Db 794 ATAAACCAAGGCAAGTGGTACTGTCTCCAAAGTCCCGGGGGAACGACGAGACCATGG 853
QY 867 ACAAAGTACTGAAAGACAAAAAAGGATAG 897
Db 854 ACAAGCCCTGGAGAAATCCAAAAAGAGAG 884

```

RESULT 13  
US-09-499-082-9  
; Sequence 9, Application us/09499082  
; Patent No. 6143522  
; GENERAL INFORMATION:  
; APPLICANT: Helbing, Caren C.  
; APPLICANT: Rabinowol, Karl  
; APPLICANT: Johnston, Randall N.  
; APPLICANT: Garkavtsev, Igor  
; TITLE OF INVENTION: METHODS OF MODULATING APOPTOSIS  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Burns, Doane, Swecker & Mathis  
; STREET: 699 Prince Street  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA



```

Qy 297 ACCAGAGAAACGCTCTACAGCAGCTTCTCCAGAGAGCAGCTAAATTAATAGTCAAGAAATTGG 356
Db 254 CGCAGAAAGCGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 313
Qy 357 GAGATGAAATTAACAGATGTTACACAAATGCTCGAATGGTGGGAAATCGGCGCAGAC 416
Db 314 GCGACGAGAAGATCCAGATCGTGAGCCAGATGGTGGAGCTGGTGGGAAACCGCAGCGGC 373
Qy 417 AAATGGAGTTACATACAGCTGTTTCCAAAGATC----- 449
Db 374 AGTGGACACCGCTGAGCTGTTCTGAGGCGCAGCAGGAGCTGGCGGACACAGTGGCA 433
Qy 450 -CTGCTGAAAGTGAACAGCCTCAGATAAAGCAAGATGGATTCCAGCCAAACAGAGAAAGA 508
Db 434 ACAGCGCAAGGTTGGCGCGGACAGGCCAATGGCGATGCGGTAGCGCAGTCTGACAAAGC 493
Qy 509 TCTTCAAGAA--GACCCCGCAGCGGACGCGGACAGGCCAATGGCGATGTTATGTACATGG 566
Db 494 CCAACAGCAGCGCTCAGCGCGGACGCGCAACAGCAGAACCGTGGAGAACCGCAGCGGC 553
Qy 567 CAATGGGATTGAAGACTGTGATGATCAGCCACCTTAAAGAAAGAAATCCCAAGTCAGCAA 626
Db 554 ACCAGACCGCAGCAGCGCGCTCGGCGCACCCAGAGAGAGAACGCCAGACCTCCA 613
Qy 627 AGAAAGAAACGCTCAAGGCCAAGCAGAGGAAAGAGCTTCACTGTTGAGTTGCAA 686
Db 614 AGAAGAAAGCGCTCCAAAGGCCAAGGCGGAGAGAGCGCTCCCTCCGACCTCCCA 673
Qy 687 TAGATCTTAATGAACCTACATACCTGCTTATGCAACCAAGTGTCTTATGTTGAGATGATAG 746
Db 674 TCGACCCCAAGCAAGTGGTACTGTCTGTCGCAACAGGCTCTCTATGGGAGATGATCG 733
Qy 747 GATGTGCAATGAACAGTGTCCAATGATGTTTCACTTTTCACTGTTTCACTTACCT 806
Db 734 GCTGCGACACGACGAGTGGCCCATCGAGTGGTTCCTCTGCTGCGTGGGCTCAATC 793
Qy 807 ATAAACCAAGGGGAATGTTATGTTCCCAAGTGCAGGAGGAGATGAATGAGAAACAATGG 866
Db 794 ATAAACCAAGGGCAAGTGGTACTGTCTCCAAAGTCCCGGGGAGAGAACGAGAACCAATGG 853
Qy 867 ACAAAGTACTGAAAGACAAAGGATAG 897
Db 854 ACAAAGCCTGGAGAAATCCAAAAAGAGAG 884

```

RESULT 15

```

; US-09-006-783A-2
; Sequence 2, Application US/09006783A
; Patent No. 6297366
; GENERAL INFORMATION:
; APPLICANT: Gudkov, Andrey V
; APPLICANT: Garbavstev, Igor
; APPLICANT: Riabowol, Karl
; TITLE OF INVENTION: p33/Ingl as a Mediator of p53 Signaling
; TITLE OF INVENTION: Pathway
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/006,783A
; FILING DATE: 15-JAN-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: No. 6297366nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 97,837
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2061 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16..897
; US-09-006-783A-2

```

```

Query Match 16.9%; Score 182.2; DB 4; Length 2061;
Best Local Similarity 56.9%; Pred. No. 3.7e-38;
Matches 393; Conservative 0; Mismatches 268; Indels 30; Gaps 2;
Qy 237 AAGAAACGTTAAAGGAAATTCATGATGCTCTACGAAATATATAAGAAAGAGATGATTTAA 296
Db 194 AACAGATCTCTGAAGCGCTAGACGAGTGTCTACGAGCGCTTCAGTCGCCGAGACAGCGGG 253
Qy 297 ACCAAGAAACGCTCTACAGCAGCTTCTCCAGAGAGCAGCACTAATTAATAGTCAAGAAATTGG 356
Db 254 CGCAGAAAGCGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 313
Qy 357 GAGATGAAATTAACAGATGTTTACAAATGCTCGAATGGTGGTGGAAATCGGCGCAGAC 416
Db 314 GCGACGAGAAGATCCAGATCTGTGAGCCAGATGGTGGAGCTGGTGGAGAACCGCAGCGGC 373
Qy 417 AAATGGAGTTACACTCACAGTGTTCCTCAAGATC----- 449
Db 374 AGTGGACAGCCAGCTGGAGCTGTTCGAGGCGCAGCAGGAGCTGGCGGACACAGTGGGCA 433
Qy 450 -CTGCTGAAAGTGAACGAGCCTCAGATAAAGCAAGATGGATTCCAGCCAAACAGAGAAAGA 508
Db 434 ACAGCGCAAGGTTGGCGCGGACAGGCCAATGGCGATGCGGTAGCGCAGTCTGACAAAGC 493
Qy 509 TCTTCAAGAA--GACCCCGCAGCGGACGCGGACAGGCCAATGGCGATGTTATGTACATGG 566
Db 494 CCAACAGCAGCGCTCAGCGCGGACGCGCAACAGCAGAACCGTGGAGAACCGCAGCGGC 553
Qy 567 CAATGGGATTGAAGACTGTGATGATCAGCCACCTTAAAGAAAGAAATCCCAAGTCAGCAA 626
Db 554 ACCAGACCGCAGCAGCGCGCTCGGCGCACCCAGAGAGAGAACGCCAGACCTCCA 613
Qy 627 AGAAAGAAACGCTCCAAAGGCCAAGCAGAGGAAAGAGCTTCACTGTTGAGTTGCAA 686
Db 614 AGAAGAAAGCGCTCCAAAGGCCAAGGCGGAGAGAGCGCTCCCTCCGACCTCCCA 673
Qy 687 TAGATCTTAATGAACCTACATACCTGCTTATGCAACCAAGTGTCTTATGGGAGATGATAG 746
Db 674 TCGACCCCAAGCAAGTGGTACTGTCTGTCGCAACAGGCTCTCTATGGGAGATGATCG 733
Qy 747 GATGTGCAATGAACAGTGTCCAATGATGTTTCACTTTTCACTGTTTCACTTACCT 806
Db 734 GCTGCGACACGACGAGTGGCCCATCGAGTGGTTCCTCTGCTGCGTGGGCTCAATC 793
Qy 807 ATAAACCAAGGGGAATGTTATGTTCCCAAGTGCAGGAGGAGATGAATGAGAAACAATGG 866
Db 794 ATAAACCAAGGGCAAGTGGTACTGTCTCCAAAGTCCCGGGGAGAGAACGAGAACCAATGG 853
Qy 867 ACAAAGTACTGAAAGACAAAGGATAG 897
Db 854 ACAAAGCCTGGAGAAATCCAAAAAGAGAG 884

```

Search completed: June 20, 2003, 02:42:11  
Job time : 85 secs



	Query Match	96.9%;	Score 1046;	DB 12;	Length 1078;
	Best Local Similarity	99.5%;	Pred. No. 2.3e-245;		
	Matches 1049; Conservative	0;	Mismatches 5;	Indels 0;	Gaps 0;
QY	1 GCGGCCGCGCCGGTGCATGTTCGGCTGCTTGAGTCGGAGCGCGCGACGCGCGCGAT 60 				
Dd	25 GC GG CG CG CG CCG GTGC ATGTGCGGCTGCTTGATCGG AGGCGCGCGCACGCGCGCAT 84 				
QY	61 CGCAGGATGTTAAGGCAGCAGCAGCAACTGTACTCTCGGCTCGGCTCTGTACC GG 120 				
Dd	85 CGCAGGATGTTAAGGCAGCAGCAGCAACTGTACTCTCGGCGCGCGCTCTGTACC GG 144 				

```
QY 121 GGAGCGAGCGGCTCCTCACCCTGCTACGTCGAGGAGTACCTTGAGTGGTGGAGTCGCT 180
Db 145 GGAGCGAGCGGCTCCTCACCCTGCTACGTCGAGGAGTACCTTGAGTGGTGGAGTCGCT 204
QY 181 GCGCCAGCATGTCAGAGGAAGTGTCTGTCTGTCGAGAGCTGGGACAAATATCAAGA 240
Db 205 GCGCCAGCATGTCAGAGGAAGTGTCTGTCTGTCGAGAGCTGGGACAAATATCAAGA 264
QY 241 AACGTTAAAGGAATGATGATGTCACGAAATAATAAGAAAGAGATGATTTAAACCA 300
Db 265 AACGTTAAAGGAATGATGATGTCACGAAATAATAAGAAAGAGATGATTTAAACCA 324
QY 301 GAAGAAACGCTACAGCAGCTTCTCCAGAGAGCAGCTAATTAATAGTCAAGAATTTGGAGA 360
Db 325 GAAGAAACGCTACAGCAGCTTCTCCAGAGAGCAGCTAATTAATAGTCAAGAATTTGGAGA 384
QY 361 TGAAGAAATACAGATTTTACACAAATGCTCGAATTTGGTGGAAATCGGGCAAGACAAT 420
Db 385 TGAAGAAATACAGATTTTACACAAATGCTCGAATTTGGTGGAAATCGGGCAAGACAAT 444
QY 421 GGAGTTTACACTCACAGTGTCTTCCAGATCCTGCTGAAAGTGAACGAGCCTCAGATAAAGC 480
Db 445 GGAGTTTACACTCACAGTGTCTTCCAGATCCTGCTGAAAGTGAACGAGCCTCAGATAAAGC 504
QY 481 AAAGATGATTCAGCCAAACCAAGAAAGATCTTCAAGAAAGACCCCGCAGCGGACCCAG 540
Db 505 AAAGATGATTCAGCCAAACCAAGAAAGATCTTCAAGAAAGACCCCGCAGCGGACCCAG 564
QY 541 TGAAGCCGCTGATTTATGTCACATGCAATGGATTTGAAGCTGTGATGATCAGCCACC 600
Db 565 TGAAGCCGCTGATTTATGTCACATGCAATGGATTTGAAGCTGTGATGATCAGCCACC 624
QY 601 TAAAGAAAGAAATCCAAAGTCAAGCAAGAAAGAAAGCAAGCTCCAGGCCCAAGCAGAAAG 660
Db 625 TAAAGAAAGAAATCCAAAGTCAAGCAAGAAAGAAAGCAAGCTCCAGGCCCAAGCAGAAAG 684
QY 661 GGAAGCTTCACTGTGTTGATTTGCAATAGATCTTATGAACCTACATATCTGTTATGCAA 720
Db 685 GGAAGCTTCACTGTGTTGATTTGCAATAGATCTTATGAACCTACATATCTGTTATGCAA 744
QY 721 CCAAGTGTCTTATGGGAGATGATAGGATGTGACATGAACAGTGTCCAAATTTGAATGGTT 780
Db 745 CCAAGTGTCTTATGGGAGATGATAGGATGTGACATGAACAGTGTCCAAATTTGAATGGTT 804
QY 781 TCACCTTTTCATGTTTTCCTACTTACCTATAAACCAGGGGAAATGGTATTGCCCAAAGTG 840
Db 805 TCACCTTTTCATGTTTTCCTACTTACCTATAAACCAGGGGAAATGGTATTGCCCAAAGTG 864
QY 841 CAGGGGAGATATGAGAAACAAATGGACAAAAGTACTGAAAGACAAAAGGATAGAAAG 900
Db 865 CAGGGGAGATATGAGAAACAAATGGACAAAAGTACTGAAAGACAAAAGGATAGAAAG 924
QY 901 ATCGAGGTAGTAAAGGCCATCCACATTTTAAAGGGTATTGCTTTTATATAATTCGTT 960
Db 925 ATCGAGGTAGTAAAGGCCATCCACATTTTAAAGGGTATTGCTTTTATATAATTCGTT 984
QY 961 TCGTTTCAGAAATGTTTATAGGTAAATGCATAAGACTATGCAATAATTTTAAATCATTA 1020
Db 985 TCGTTTCAGAAATGTTTATAGGTAAATGCATAAGACTATGCAATAATTTTAAATCATTA 1044
QY 1021 GTATTAATGGTATTAAGAAAGTGTGTACTTTG 1054
Db 1045 GTATTAATGGTATTAAGAAAGTGTGTACTTTG 1078
```

RESULT 2  
US-10-115-899-6  
; Sequence 6, Application US/10115899  
; Patent No. US20020151025A1  
; GENERAL INFORMATION:  
; APPLICANT: Otsuka Pharmaceutical Co., Ltd.  
; TITLE OF INVENTION: Human TSC403 gene and human ING1L gene  
; FILE REFERENCE: Q60193

```
; CURRENT APPLICATION NUMBER: US/10/115,899  
; CURRENT FILING DATE: 2002-04-05  
; PRIOR APPLICATION NUMBER: 09/601,478  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: JP H10-134679  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: JP H10-73234  
; PRIOR FILING DATE: 1998-03-05  
; PRIOR APPLICATION NUMBER: JP H10-38133  
; PRIOR FILING DATE: 1998-02-03  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 840  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: human embryonic brain cDNA library  
US-10-115-899-6  
  
Query Match 77.6%; Score 838.4; DB 12; Length 840;  
Best Local Similarity 99.9%; Pred. No. 1e-194; 1; Indels 0; Gaps 0;  
Matches 839; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 68 ATGTTAGGCGCAGCAGCAGCAACTGTACTCTCGGCTCGCTCTCGTACCGGGGAGCGG 127  
Db 1 ATGTTAGGCGCAGCAGCAGCAACTGTACTCTCGGCTCGCTCTCGTACCGGGGAGCGG 60  
QY 128 AGCCGGCTGCTCACCTGCTACGTGAGGACTACCTTGAGTGGTGGAGTCGTCGCCAC 187  
Db 61 AGCCGGCTGCTCACCTGCTACGTGAGGACTACCTTGAGTGGTGGAGTCGTCGCCAC 120  
QY 188 GACATGCGAGCAAGCTGTCTGTGTCGAGAGCTGCGACAAACAAATATCAAGAAAGCTTA 247  
Db 121 GACATGCGAGCAAGCTGTCTGTGTCGAGAGCTGCGACAAACAAATATCAAGAAAGCTTA 180  
QY 248 AAGGAAATGTGATGTCTTACGAAATAATAAGAAAGAGATGATTTAAACCAAGAA 307  
Db 181 AAGGAAATGTGATGTCTTACGAAATAATAAGAAAGAGATGATTTAAACCAAGAA 240  
QY 308 CGCTACAGCAGCTTCTCCAGAGAGCAGCTAATTAATAGTCAAGAAATGGAGATGAA 367  
Db 241 CGCTACAGCAGCTTCTCCAGAGAGCAGCTAATTAATAGTCAAGAAATGGAGATGAA 300  
QY 368 ATACAGATGTTTACACAAATGCTCGAATTTGGTGGAAATCGGCAAGCAAAATGGAGTTA 427  
Db 301 ATACAGATGTTTACACAAATGCTCGAATTTGGTGGAAATCGGCAAGCAAAATGGAGTTA 360  
QY 428 CACTCACAGTGTTCACAGATCCTGCTGAAAGTGAACGAGCCTCAGATTAAGCAAGATG 487  
Db 361 CACTCACAGTGTTCACAGATCCTGCTGAAAGTGAACGAGCCTCAGATTAAGCAAGATG 420  
QY 488 GATTCAGCCACCAAGAAAGATCTTCAAGAAAGCCCGCAGCGGACCAAGTGAAGC 547  
Db 421 GATTCAGCCACCAAGAAAGATCTTCAAGAAAGCCCGCAGCGGACCAAGTGAAGC 480  
QY 548 CGTGATTTATGTCACATGCGCAATGGATTGAAGCTGTGATGATCAGCCACCTTAAAGAA 607  
Db 481 CGTGATTTATGTCACATGCGCAATGGATTGAAGCTGTGATGATCAGCCACCTTAAAGAA 540  
QY 608 AAGAAATCCAGTCAGCAAGAAAGAAAGAAAGCCTTCAAGCCCAAGCAAGGAGGAGCT 667  
Db 541 AAGAAATCCAGTCAGCAAGAAAGAAAGAAAGCCTTCAAGCCCAAGCAAGGAGGAGCT 600  
QY 668 TCACCTGTTGAGTTTGCATAGATCTTAATGAACCTACATCTGCTTATGCAACCAAGTG 727  
Db 601 TCACCTGTTGAGTTTGCATAGATCTTAATGAACCTACATCTGCTTATGCAACCAAGTG 660  
QY 728 TCTTATGGGAGATGATAGGATGTGACATGAACAGTGTCCAAATTTGATTTTCACTTT 787  
Db 661 TCTTATGGGAGATGATAGGATGTGACATGAACAGTGTCCAAATTTGATTTTCACTTT 720  
QY 788 TCATGTTTCTACTTACCTATATAACCAAGGGAATGTTATGCCCAAGTGCAGGGA 847
```





Db	408	CAATGGCGATGCGGTAGCGCAGTCTGACAAAGCCCAACAGCAAGCGCTCACGGGGCAGCG	467
Qy	535	GACCAGTGAAGCCGCTGATTATGTCTCACATGGCAAAATGGGATTGAAGACTGTGATGATCA	594
Db	468	CAACAACGAGAACCGTGAGAACCGTCCACGCAACCCAGCACCGCGCCCTCGGG	527
Qy	595	GCCACCTTAAAGAAAGAAATCCAAGTCAGCAAGAAAAAAGAAACGCTTCCAAGGCCAAGCA	654
Db	528	CACACCCCAAGGAGAAAGGCCAAGACCTCCAAGAAGAAGAGCGTCCAAGGCCAAGGC	587
Qy	655	GGAAGGGAGCTTCACCTGTTGAGTTTGCAATAGATCCTTAATGAACCTACATACTGCTT	714
Db	588	GAGAGCAGAGGCGTCCCTGCCACCTCCCCATCGAGCCCCCAAGCAACCCACGTACTGTCT	647
Qy	715	ATCGAACCAAGTGTCTTTATGGGAGATGATAGGATGTGACAATGAACAGTGTCCAATTGA	774
Db	648	GTGCAACCAAGTCTCCTATGGGAGATGATCGGCTGGGACAACGACGAGTGCCCCATCGA	707
Qy	775	ATGGTTTCAGTTTTCATGTGTTTTCATCTTACTTATAACCAAGGGGAAATGGTATTGCCC	834
Db	708	GTGGTTCACCTTCTCGTGCGTGGGGCTCAATCATATAACCCCAAGGGCAAGTGTACTGCC	767
Qy	835	AAAGTCAGGGGAGATTAATCAGAAAAACAATGGCAAAAAGTACTTCAAAGACAAAAAGGA	894
Db	768	CAAGTCGGGGGGAGACCGAGAGACCATGGCAAAAGCCCTGGAGAAATCCAAAAAAGA	827
Qy	895	TAG 897	
Db	828	GAG 830	

## RESULT 5

US-09-764-877-3454  
; Sequence 3454, Application US/09764877  
; Patent No. US20020147140A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC005  
; CURRENT APPLICATION NUMBER: US/09/764, 877  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 4031  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3454  
; LENGTH: 8487  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-877-3454

	Qy	505	AAGATCTTCAAGAAGACCCCGCAGGCGACCGAGTCGAATTTATGTGCACAT	564
	Dd	6529	CCCAAACAGCAAG- CGGTACGGGGCAGCGCAACAACGAGAACC GTGAGAACGCGTCCAG	6587
	Qy	565	GGCAARTGGGATTGAAGACTGTGATGATCAGCCACCTTAAGAAAGAAATAATCCAAGTCAGC	624
	Dd	6588	CAACCAGCACCGACGACGGCGGCCCTGGGCACACCCCAAGGAGAAAGGCCAAGACCTC	6647
	Qy	625	AAAGAAAAAAGAACGGCTCCAAGGCGCAAGCAAGGAAAGGGAAGCTTCACCTGTTCAGGTTTTGC	684
	Dd	6648	CAGAGACAAGACGGCTCCAAGGCGCAAGCGGAGCGAGGCGTCCGCTGCCACCTCCC	6707
	Qy	685	AATAGATCCPAATGAACCTACATCTACTGCTTTATGCAACCAAGTGTCTTATGGGAGATGAT	744
	Dd	6708	CATCGACCCCAACGAACCCACGCTACTGTCTGTGCAACCAAGTGCTCCTATGGGAGATGAT	6767
	Qy	745	AGGATGTGACAAATGAACAGTGTCCAATTGAATGGTTTCAGTTCATGTGTTTCATTAC	804
	Dd	6768	CGGCTGGGACACGACGAGTGGCCCCCATCGAGTGGTTCACCTCTCGTGGGTGGGGCTCAA	6827
	Qy	805	CTATAAACCAAGGGGAAGATGGTATTGCCCAAAGTGCAGGGGGAGATAATGAGAAAAACAAT	864
	Dd	6828	TCATAAACCAAGGGCAAGTGGTACTGTGCCAAGTGCCGGGGGAGACGAGAGACCAT	6887
	Qy	865	GGACAAAAGTACTGAAAAAGACAAAAAAGGATAG	897
	Dd	6888	GGACAAGGCCCTGGAGAAATCCAAAAGAGAG	6920

## RESULT 6

US-09-968-653A-2  
Sequence 2, Application US/09968653A  
Publication No. US20030073084A1  
GENERAL INFORMATION:  
APPLICANT: Gudkov, Andrey V  
Garkavstev, Igor  
Rindowol, Karl  
TITLE OF INVENTION: p33INGl as a Mediator of p53 Signaling  
Pathway  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
STREET: 300 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/968,653A  
FILING DATE: 01-Oct-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/006,783A  
FILING DATE: 15-JAN-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: No. US20030073084Alnan, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 97,837  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-913-0001  
TELEFAX: 312-913-0002  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2061 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear



```

: Publication No. US20030104529A1
:
: GENERAL INFORMATION:
: APPLICANT: Tang, Y. Tom
: APPLICANT: Liu, Chenghua
: APPLICANT: Asundi, Vinod
: APPLICANT: Zhang, Jie
: APPLICANT: Ren, Felyan
: APPLICANT: Chen, Rui-hong
: APPLICANT: Zhao, Qing A.
: APPLICANT: Wehrman, Tom
: APPLICANT: Xue, Aildong J.
: APPLICANT: Yang, Yonghong
: APPLICANT: Wang, Jian-Rui
: APPLICANT: Zhou, Ping
: APPLICANT: Ma, Yunding
: APPLICANT: Wang, Dunrui
: APPLICANT: Wang, Zhwei
: APPLICANT: Tillinghast, John
: APPLICANT: Drmanac, Radoje T.
:
: TITLE OF INVENTION: No. US20030104529A1: Polypeptide
: FILE REFERENCE: 784CIP2B
:
: CURRENT APPLICATION NUMBER: US20030104529A1
: CURRENT FILING DATE: 2002-01-11
: PRIOR APPLICATION NUMBER: 09/550-425
: PRIOR FILING DATE: 2000-04-25
: PRIOR APPLICATION NUMBER: 09/440-121
: PRIOR FILING DATE: 2000-01-21
: NUMBER OF SEQ ID NOS: 1104
: SOFTWARE: pt_FL-genes Version 1.0
: SEQ ID NO 435
:
: LENGTH: 1864
:
: TYPE: DNA
: ORGANISM: Homo sapiens
:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (164)..(1420)
: US-10-037-270-435

```

Query Match	8.5%	Score	91.6;	DB	9;	Length	1864;
Best Local Similarity	64.8%;	Pred. No.	4.1e-12;				
Matches	136;	Conservative	0;	Mismatches	74;	Indels	0;
Gaps	0;						
QY	674	GTTGAGTTTGCAATAGATCCTAATGAACCTACATACTGCTTTATGCAACCAAGTGTCTTAT	733				
Db	1214	GTTGATTGGACITTCAGACCCAATGAACCTCGATATGCAATTTGTAATCAGGATATCTTAT	1273				
QY	734	GGGGAGATGATGAGTGTGACAAATGAACAGTGTCCAAATGAAATGGTTTTCACITTTTCATGT	793				
Db	1274	GGTGAGATGGTGGGATGTGATACCAAGATGGCCCTATAGAATGGTTTCCATTATGGCTGC	1333				
QY	794	GTTTTCACITTCACCTATAAAACCAAGGGGAAATGCTATGGCCCAAAGTGCAGGGGAGATAAT	853				
Db	1334	GTTCGATTGCAGAGGCCCAAAAGGCAATGTGCTGTCCACAGTGCACTGTGCAATG	1393				
QY	854	GAGAAACAATGGACAAAAGTACTGAAAAAG	883				
Db	1394	AAGAGAAGAGGCGACGACACAAATAAAGG	1423				

RESULT 9  
US-09-854-133-146  
; Sequence 146, Application US/09854133  
; Publication No. US20020183499A1  
; GENERAL INFORMATION:  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Mohamath, Raodoh  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Secrist, Heather  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.475C10

```

; CURRENT APPLICATION NUMBER: US/09/854.133
;
; CURRENT FILING DATE: 2001-05-11
;
; NUMBER OF SEQ ID NOS: 735
;
; SOFTWARE: FastSEQ for Windows Version 3.0
;
; SEQ ID NO 146
;
; LENGTH: 451
;
; TYPE: DNA
;
; ORGANISM: Homo sapien
US-09-854-133-146

```

	Query Match	7.7%	Score 83.6;	DB 9;	Length 451;
	Best Local Similarity	58.6%;	Pred. No. 1.5e-10;		
	Matches 163;	Conservative 0;	Mismatches 110;	Indels 1;	Gaps 1;
Qy	169	CGTGGAGTCTGCTGCCCCACACACATGCAGAGGAAGCTGTCTGTGCTGCGAGAGCTGGACAA	228		
Db	16	CATCGAGTCCCTGCTTTTCGACTTCGACATTCGACAGAAATGTCTCGCTGATCGGGAGATCGAGC	75		
Qy	229	CAAAATACAGAAACGTTAAAGGAAATTTGATGATGTCTACGAAAAATATAAAGAAAGA	288		
Db	76	GAATACCAAGAGATCTCTGAAGGAGCTAGACGAGTGTACGAGCGCTTCAGTCGCGAGAC	135		
Qy	289	TGATTTTAAACCAAGAAGACGCTCTACAGCAGCTTCTCCAGAGAGACACTAATTAATAGTCA	348		
Db	136	AGACGGGGCGCAGAAGCGCGGATGCTGCACTGTGTGCAGCGCGCTGATCCGCA - CCA	194		
Qy	349	AGATTGGGAGATGAAAAAATACAGATGTTACACAAATGCTCAAAATGGTGGAAATCG	408		
Db	195	GGAGCTGGGGCAGAGAAGATCCAGATCGTGAGCCAGATGGTGAGCTGGTGGGAACCG	254		
Qy	409	GGCAAGACAATGGAGTTTACACTCACAGTGTTCCTCAAG	446		
Db	255	CACGCGCAGGTGGACAGCCACGCTGGAGCTGTTCGAGG	292		

```

RESULT 10
US-09-738-973-146
; Sequence 146, Application US/09738973
; Patent NO. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indirias, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliot, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09/738,973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 146
; LENGTH: 451
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-738-973-146

```

	Query Match	7.7%	Score 83.6	DB 10	Length 451
	Best Local Similarity	58.6%	Pred. No. 1.5e-10		
	Matches 163	Conservative	0	Mismatches 114	Indels 1
	Gaps				
Qy	169	CGTGGAGTCTGCTGCCCCACACATGCAGAGGAAGCTGCTCTGCTGCGAGACGCTGGACAA	228		
Db	16	CATCTGAGTCCCTGCTTTTCGACTTGCAGAGAAATGTCTCGCTGATCGGGAGATCGACGC	75		



```
Db 67 GCACTATCGGACGATATCGAAGCCTTCCCTCGGAACCTTCAGAGAACTTCAGCTGAT 126
QY 214 GCGAGAGCTGGACAAACAAATATCAAGAAACGTTAAAGAAATGATGCTCTACGAAAA 273
Db 127 GCGAGAGCTGGACGAGGACGAAATGAAGAGCAGAGATGACATCTGGCTGCAGA 186
QY 274 ATATAA-----GAAAGAGATGATTTAAACCAAGAAACGCTTACAGCAGCTTCT 324
Db 187 GTACATCTCCACGGTGAAGACGCTGTCTCCAGACACCGCGCTGGAGCGCTGCAGAAGAT 246
QY 325 CCAGAGACGACATAATTAATAGTCAAGAATGGAGATGAAAAATACAGATGTTTACACA 384
Db 247 CCAGAACGCTTACAGCAATGGAAGGAATACAGTGACGACAAAGTGCAGCTGGCCATGCA 306
QY 385 AATGCTCGAATGGTGGAAATCGGCAAGACAAATGGAGTTACACTCACAGTGTTCACA 444
Db 307 GACCTACGAGATGGTGGTAACACACATTCGAAGGCTTGATGCAGACCTGGCGGCTTTGA 366
QY 445 AG 446
Db 367 AG 368
```

## RESULT 14

```
US-09-960-352-4582/c
; Sequence 4582, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warrier, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 4582
; LENGTH: 393
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 20-LIB3057-013-Q1-K1-E7
US-09-960-352-4582
```

```
Query Match 4.8%; Score 51.6; DB 10; Length 393;
Best Local Similarity 51.3%; Pred. No. 0.0089;
Matches 120; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 847 AGATATGAGAAACAAATCGACAAAGTACTGAAAGACAAAAAGGATAGAGATCGAG 906
Db 387 AAAAAATAAAAAATTAATTTAAAAATTAATAAAAAAAATTTATAAAAAATTTA 328
QY 907 GTAGTAAGGCCATCCACATTTTAAAGGTTATTGCTCTTTTATATAATTCGTTGCTTT 966
Db 327 AATTTATAATAATAATAATAATATCAATAATAATAATAATAATAATAATAATAATA 268
QY 967 CAGAAAAATGTTTAGGGTAAATGCAAGACTATGCAATAATTTTAAATCATTAGTATTA 1026
Db 267 TAAAAATATATATTTAAAAAATAAAAAAATTTTAAATAATTTTAAAAATAAAA 208
QY 1027 ATGGTGATTAATAAGTGTGTGACTTTGAAAAAATAAAAAAATAAAAAAATAAAAA 1080
Db 207 AAATTATAAATAATAATAATAATAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 154
```

## RESULT 15

```
US-09-919-580-114/c
; Sequence 114, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
```

```
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 114
; LENGTH: 361
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc.feature
; LOCATION: 49, 89, 101, 141, 209, 253, 305, 324, 331, 360
; OTHER INFORMATION: n = A,T,C or G
US-09-919-580-114

Query Match 4.7%; Score 50.8; DB 10; Length 361;
Best Local Similarity 64.6%; Pred. No. 0.013;
Matches 73; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 968 AGAAATGTTTGGTAAATGCATGAAGACTATGCAATATTTTAAATCATTAGTATTA 1027
Db 124 ATAAAAAGTATTTTGGTCACTGTNTTACATGTTAANAGGGACTGAAATAGTTATATTA 65
QY 1028 TGGTGATTTAAAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1080
Db 64 GTTGTGATTAATAATTTTAAATAATAATAATAATAATAATAATAATAATAATAATA 12
```

Search completed: June 20, 2003, 02:45:35  
Job time : 195 secs

Result	Query No.	Score	Match	Length	DB	ID	Description
	1	812.8	75.3	1464	11	AK012716	Mus musc
	2	747.8	69.2	912	14	BQ377444	Mus musc
c	3	735	68.1	735	14	BM982877	UI-CF-EN1
c	4	727.6	67.4	798	12	BG184056	RST2972 A
	5	711	65.8	793	12	BE796780	BE796780
	6	694.2	64.3	778	13	BI548536	BI548536
	7					603191255	603191255

## ALIGNMENTS

<b>TITLE</b>	Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
<b>JOURNAL</b>	Genome Res. 10 (10), 1617-1630 (2000)
<b>MEDLINE</b>	20499374
<b>PUBMED</b>	11042159





```

Qy 536 ACCAGTGAAGCCGCTGATTATGTCACATGCGCAATGGGATGTAAGACTGTGATGATCAG 595
Db 948 ACCAGTGAAGCCGCTGATTATGTCACATGCGCAATGGGATGTAAGACTGTGATGATCAG 1007
Qy 596 CCACCTTAAGAAAAGAAATCCAACTCAGCAAGAAAAGAAAGAAAGAAAGAAAGAAAGAAAG 655
Db 1008 CCACCGAAGAAAAGAAATCCAACTCAGCAAGAAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1067
Qy 656 GAAGAGGAAGCTTCCACCTGTTGAGTTTGCATATAGATCCTATGAAACCTACATATGCTTA 715
Db 1068 GAGAGGAGGAGCATCCCTGTCGAGTTTGCCATGATCCCATGAGCCACCTACTGCTTG 1127
Qy 716 TGCACCAAGAGTCTTATGGGAGATGATGAGGATGACAAATGAACAGTGTCCAAATGAA 775
Db 1128 TGTAAACCAAGTGTCTTACGGGAGATGATAGGCTGTGACAAATGACAACTGTCCCAATGAA 1187
Qy 776 TGGTTTCACCTTTTCATGTTTCACTTACCTATATAAACCAAGGGAATGTAATGCCCCA 835
Db 1188 TGGTTTCACCTTTTCATGTTTCACTTACCTATATAAACCAAGGGAATGTAATGCCCCA 1247
Qy 836 AAGTGCAGGGGAGATATAGAAAACAAATGACAAAAGTACTGAAAAGACAAAAAAGGAT 895
Db 1248 AAGTGTAGGGGAGACATGAGAAAACCAATGACAAAAGTACTGAAAAGACAAAAAAGGAG 1307
Qy 896 AGAAGATCGAGGTAGTAAAGCCATCCACATTTTAAAGGGTATTTGCTTTTATATAAT 955
Db 1308 AGAAGAGCGAGGTAGTAAAGCCATC--GGTTTAAAGGGTCTTTGCTTTTATAGAA 1365
Qy 956 TCGTTTCTTTCAGAAAATGTTTAAAGGTAATGCATAAGACTATGCAATATTTTAAAT 1015
Db 1366 TCGTTTCTTTCAGAAAATGTTTAAAGGTAATGCATAAGACTATGCAATATTTTAAAT 1425
Qy 1016 CATTAGTATTATGTTGTTATTAAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1054
Db 1426 CATTAGTATTGTTGTTGTTATTAAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1464

RESULT 2
BQ277444
LOCUS AGENCOURT_6738254 NIH_MGC_127 Homo sapiens CDNA clone IMAGE:5810745
DEFINITION 5', mRNA sequence.
ACCESSION BQ277444
VERSION BQ277444.1 GI:20487652
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nhl.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: NCI
CDNA Library Preparation: Michael Brownstein Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2058 row: h column: 10
High quality sequence stop: 587.
Location/Qualifiers
1..912
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5810745"
/clone_lib="NIH_MGC_127"
/tissue_type="mixed (pool of 40 RNAs)"
/lab_host="DH10B (T1-phage-resistant)"
/note="Vector: pDNR-LIB; Site_1: SfiI (ggccattatggcc);

```

```

Site_2: SfiI (ggccgctcgcc); Double-stranded cDNA was
prepared from a pool of 40 cell line polyA+ RNAs (bladder
- 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -
4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,
kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%,
ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary
gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were
used in cloning as follows:
5'-AAGCAGTGTGATCAACGCGGCGGCACATG-dt(30)NN-3'. Full-length
enriched library was constructed using the Clontech
Creator SMART kit and size-selected to contain the 1-2 kb
size fraction (other fractions present in NIH_MGC_126 and
NIH_MGC_128). Library created in the laboratory of T.
Usdin, M.D., Ph.D. (NIMH, NIH). Note: this is a NIH_MGC
Library."
BASE COUNT 313 a 184 c 235 g 180 t
ORIGIN
Query Match 69.2%; Score 747.8; DB 14; Length 912;
Best Local Similarity 96.9%; Pred. No. 2.5e-104;
Matches 794; Conservative 0; Mismatches 22; Indels 3; Gaps 3;
Qy 47 GCGAGCGCGGATCGGCAGGATGTTAGGCGAGCAGCAGCAGCACTACTCTCGTGGCT 106
Db 1 GCGAGCGCGGATCGGCAGGATGTTAGGCGAGCAGCAGCAGCACTACTCTCGTGGCT 60
Qy 107 GCGCTCTCTGCGGGGAGCGGCTGCTCCTCTGCTCAGTCTCAGGAGCTACCTTCAG 166
Db 61 GCGCTCTCTGCGGGGAGCGGCTGCTCCTCTGCTCAGTCTCAGGAGCTACCTTCAG 120
Qy 167 TCGCTGGAGTCTGCTGCCCCACGATGCGAGGAGCACTGCTGCTGCTCGAGAGCTGGAC 226
Db 121 TCGCTGGAGTCTGCTGCCCCACGATGCGAGGAGCACTGCTGCTGCTCGAGAGCTGGAC 180
Qy 227 AACAAATATCAAGAAACGTTAAAGAAATGATGATGCTCAGAAAAATATAAGAAAGA 286
Db 181 AACAAATATCAAGAAACGTTAAAGAAATGATGATGCTCAGAAAAATATAAGAAAGA 240
Qy 287 GATGATTTAAACCAAGAAACGTTACAGAGCTTCTCCAGAGAGCACTAATTAATAGT 346
Db 241 GATGATTTAAACCAAGAAACGTTACAGAGCTTCTCCAGAGAGCACTAATTAATAGT 300
Qy 347 CAAGAATTTGGAGATGAAAAAATACAGATTTTACACAAATGCTCGAATTTGGTGAAT 406
Db 301 CAAGAATTTGGAGATGAAAAAATACAGATTTTACACAAATGCTCGAATTTGGTGAAT 360
Qy 407 CGGCGAGCAAAATGGAGTTACATCAGTGTTCCTCAAGATCTCTGCTGAAGTGAAGA 466
Db 361 CGGCGAGCAAAATGGAGTTACATCAGTGTTCCTCAAGATCTCTGCTGAAGTGAAGA 420
Qy 467 GCCTCAGATAAGCAAGATGATGATTCAGCAACCAAGAAAGATCTTCAAGAGAGCCCGC 526
Db 421 GCCTCAGATAAGCAAGATGATGATTCAGCAACCAAGAAAGATCTTCAAGAGAGCCCGC 480
Qy 527 AGGCGAGCGGACAGTGAAGCGGCTGATTTATGTCACATGGCAATGGGATTTGAAGCTGT 586
Db 481 AGGCGAGCGGACAGTGAAGCGGCTGATTTATGTCACATGGCAATGGGATTTGAAGCTGT 540
Qy 587 GATGATCAGCCACCTTAAGAAAGAAATCCAAAGTCAAGAAAGAAAGAAAGAAAGCGTCCAAG 646
Db 541 GATGATCAGCCACCTTAAGAAAGAAATCCAAAGTCAAGAAAGAAAGAAAGAAAGCGTCCAAG 600
Qy 647 GCCAAGCAGGAAAGGAGGCTTCACCTGCTGAGTTTGCATATGATGATGATGATGATGATGAT 706
Db 601 GCCAAGCAGGAAAGGAGGCTTCACCTGCTGAGTTTGCATATGATGATGATGATGATGATGAT 660
Qy 707 TACTGCTTATGCAACCAAGTGTCTTATGGGAGATGATGATGATGATGATGATGATGATGAT 766
Db 661 TACTGCTTATGCAACCAAGTGTCTTATGGGAGATGATGATGATGATGATGATGATGATGAT 719
Qy 767 CCAATTGAATGGTTTAC-TTTTATGATGTTTCTTACCTTACCTTACCTTACCTTACCTTACCT 824
Db 767 CCAATTGAATGGTTTAC-TTTTATGATGTTTCTTACCTTACCTTACCTTACCTTACCTTACCT

```

```

Db      720 CCAATTGAATGGTTCACATTTTTCATGCTGGTTCACCTAAGTATATAAAACCAAGGGGAAAT 779
QY      825 GGTATTGCCCAAGTGCAGGGGAGATAATGAGAAACAA 863
      ||||| ||| | ||| ||| | ||| ||| |
Db      780 GGTATTGCCACAGTGCAGGGGAGATATGGAGAAA 818

RESULT 3
BM982877/c
LOCUS      735 bp mRNA linear EST 21-MAR-2002
DEFINITION UI-CF-EN1-acs-d-05-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone
ACCESSION  BM982877
VERSION     UI-CF-EN1-acs-d-05-0-UI 3', mRNA sequence.
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
REFERENCE   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS     Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE       Normalization and subtraction: two approaches to facilitate gene
            discovery
JOURNAL     Genome Res. 6 (9), 791-806 (1996)
MEDLINE     9704477
COMMENT     Contact: McCray, PB
            University of Iowa
            2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
            Tel: 319 356 4866
            Fax: 319 356 7171
            Email: paul-mccray@uiowa.edu
            Tissue Procurement: Dr. M. J. Welsh, University of Iowa
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 FORWARD
            POLYA=Yes.

FEATURES             Location/Qualifiers
     1..735
     /organism="Homo sapiens"
     /db_xref="taxon:9606"
     /clone="UI-CF-EN1-acs-d-05-0-UI"
     /clone_lib="UI-CF-EN1"
     /tissue_type="Primary Lung Cystic Fibrosis Epithelial
     Cells"
     /dev_stage="Adult"
     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
     /note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
     modified polylinker; Site.1: EcoR I; Site.2: Not I;
     UI-CF-EN1 is a normalized cDNA library containing the
     following tissue(s): Primary Lung Cystic Fibrosis
     Epithelial Cells. The library was constructed according to
     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
     1996. First strand cDNA synthesis was primed with an
     oligo-dT primer containing a Not I site. Double stranded
     cDNA was ligated to an EcoR I adaptor, digested with Not
     I, and cloned directionally into pT73-Pac vector. The
     oligonucleotide used to prime the synthesis of
     first-strand cDNA contains a library tag sequence that is
     located between the Not I site and the (dT)18 tail. The
     sequence tag for this library is CTGCTCAGGT.
     TAG_LIB=UI-CF-EN1
     TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS
     6hr to LPS 24h
     TAG_SEQ=CTGCTCAGGT"
BASE COUNT      180 a 161 c 114 g 280 t
ORIGIN
Query Match      68.1%; Score 735; DB 14; Length 735;
Best Local Similarity 100.0%; Pred. No. 2.4e-102;

```

```

Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      338 ATTAATAGTCAAGAAATGGGAGATGAAATAATACAGATTCTTACACAAATGCTCGAATTG 397
      ||||| ||| | ||| ||| | ||| ||| |
Db      735 ATTAATAGTCAAGAAATGGGAGATGAAATAATACAGATTCTTACACAAATGCTCGAATTG 676
QY      398 GTGAAAAATCGGGCAAGACAAATGGAGTTTACACTCACAGTGTGTTTCCAAAGATCCTGCTGAA 457
      ||||| ||| | ||| ||| | ||| ||| |
Db      675 GTGAAAAATCGGGCAAGACAAATGGAGTTTACACTCACAGTGTGTTTCCAAAGATCCTGCTGAA 616
QY      458 AGTGAACGAGCCTCAGATAAAGCAAGATGGATTCAGCCAGCAACCAAGAGATCTTCAAGA 517
      ||||| ||| | ||| ||| | ||| ||| |
Db      615 AGTGAACGAGCCTCAGATAAAGCAAGATGGATTCAGCCAGCAACCAAGAGATCTTCAAGA 556
QY      518 AGACCCCGCAGCGCAGCGGACCTGTAAGCCGCTGATTATGTGCACATGGCAATGGGATT 577
      ||||| ||| | ||| ||| | ||| ||| |
Db      555 AGACCCCGCAGCGGAGCGGACCTGTAAGCCGCTGATTATGTGCACATGGCAATGGGATT 496
QY      578 GAAGACTGTGTGATGATCAGCCACCTAAAGAAAAAGAAATCCAAAGTCAGCAAGAAAAAGAAA 637
      ||||| ||| | ||| ||| | ||| ||| |
Db      495 GAAGACTGTGTGATGATCAGCCACCTAAAGAAAAAGAAATCCAAAGTCAGCAAGAAAAAGAAA 436
QY      638 CGCTCAAGGCCAAGCAGGAAGGAGCTTCACCTGTTGAGTTTCAATAGATCTTAAT 697
      ||||| ||| | ||| ||| | ||| ||| |
Db      435 CGCTCAAGGCCAAGCAGGAAGGAGCTTCACCTGTTGAGTTTCAATAGATCTTAAT 376
QY      698 GAACCTACATACCTGCTTATGCAACCAAGTGTCTTATGGGAGAGATAGGATGTGACAAT 757
      ||||| ||| | ||| ||| | ||| ||| |
Db      375 GAACCTACATACCTGCTTATGCAACCAAGTGTCTTATGGGAGAGATAGGATGTGACAAT 316
QY      758 GAACAGTGTCCAAATGAATGGTTTCACTTTTCAATGTTTCACTTACCTATACCAACCAAG 817
      ||||| ||| | ||| ||| | ||| ||| |
Db      315 GAACAGTGTCCAAATGAATGGTTTCACTTTTCAATGTTTCACTTACCTATACCAACCAAG 256
QY      818 GGAATGTGATTTGCCAAGTCGAGGGAGATATGAGAAAAACAATGGAACCAAGTACT 877
      ||||| ||| | ||| ||| | ||| ||| |
Db      255 GGAATGTGATTTGCCAAGTCGAGGGAGATATGAGAAAAACAATGGAACCAAGTACT 196
QY      878 GAAAGACAAAAAAGGATAGAGATCGAGGTAGTAAAGGCCATCCACATTTTAAAGGGTT 937
      ||||| ||| | ||| ||| | ||| ||| |
Db      195 GAAAGACAAAAAAGGATAGAGATCGAGGTAGTAAAGGCCATCCACATTTTAAAGGGTT 136
QY      938 ATTTGCTTTTATATAATTCGTTGCTTTCAGAAAAATGTTTATAGGTAATGATGATAGAC 997
      ||||| ||| | ||| ||| | ||| ||| |
Db      135 ATTTGCTTTTATATAATTCGTTGCTTTCAGAAAAATGTTTATAGGTAATGATGATAGAC 76
QY      998 TAGCAATATTTTAAATCATTAGTATTATGTTGTTATTAAGAGTTGTTGTTGTTGTTGTTG 1057
      ||||| ||| | ||| ||| | ||| ||| |
Db      75 TATGCAATATTTTAAATCATTAGTATTATGTTGTTTAAAGAGTTGTTGTTGTTGTTGTTG 16
QY      1058 AAAAAAAGAAAAA 1072
      ||||| ||| | ||| ||| |
Db      15 AAAAAAAGAAAAA 1
      ||||| ||| | ||| ||| |

RESULT 4
BM984056/c
LOCUS      798 bp mRNA linear EST 21-APR-2001
DEFINITION RST2972 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION  BM984056
VERSION     BM984056.1 GI:13705743
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
REFERENCE   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1. (bases 1 to 798)
AUTHORS     Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
            Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.,
            Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith
            ,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
            ,J., Danzig,J. and Ducar,M.
            Creation of genome-wide protein expression libraries using random

```

activation of gene expression  
 Nat. Biotechnol. 19 (5), 440-445 (2001)  
 21227151  
 COMMENT Contact: Scott J. Cain  
 Athersys, Inc.  
 3201 Carnegie Ave, Cleveland, OH 44115, USA  
 Tel: 216 431 9900  
 Fax: 216 361 9596  
 Email: scai@atersys.com  
 High quality sequence stop: 531.  
 Location/Qualifiers

# FEATURES

source

1. 798  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="Athersys RAGE Library"  
 /cell\_line="HT1080"  
 /note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."  
 BASE COUNT 197 a 178 c 130 g 293 t

ORIGIN  
 Query Match 67.4%; Score 727.6; DB 12; Length 798;  
 Best Local Similarity 98.9%; Pred. No. 3.1e-101;  
 Matches 785; Conservative 0; Mismatches 4; Indels 5; Gaps 5;

261 ATGCTACGAAAATATAGAAAGAGATGATTTAAACCCAGAGAACGGTCTACAGCAGC 320  
 Db ATGCTACGAAAATAT-AGAAAGAGATGATTTAAACCCAGAGAACGGTCTACAGCAGC 731  
 QY 321 TTCTCCAGAGAGCAGCACTAATTAATAGTCAAGAAATGGGAGATGAAGAAATACAGATTGTTA 380  
 Db 730 TTCTCCAGAGAGCAGC-ATTATAGTCAAGAAATGGGAGATGAAGAAATACAGATTGTT- 674  
 QY 381 CACAATATGCTCGAATTTGGTGGAAATCGGGCAAGACAAATGGAGTTACACTACAGTGT 440  
 Db 673 CACAATATGCTCGAATTTGGTGG- AAATCGGGCAAGACAAATGGAGTTACACTACAGTGT 615  
 QY 441 TCCAAGATCTGCTGAAGTGAAGAGAGCCTCAGATTAAGCAAGATGGATTCAGGCAAC 500  
 Db 614 TCCAAGATCTGCTGAAGTGAAGAGAGCCTCAGATTAAGCAAGATGGATTCAGGCAAC 555  
 QY 501 CAGAAAGATCTTCAAGAGACCCCGCAGGCGAGCCAGTGAAGCCGTGATTTATGTC 560  
 Db 554 CAGAAAGATCTTCAAGAGACCCCGCAGGCGAGCCAGTGAAGCCGTGATTTATGTC 495  
 QY 561 ACATGGCAATGGGATTTGAAGACTGTGATGATCAGCCACCTAAAGAAAAGAAATCCAAAGT 620  
 Db 494 ACATGGCAATGGGATTTGAAGACTGTGATGATCAGCCACCTAAAGAAAAGAAATCCAAAGT 435  
 QY 621 CAGCAAGAAAAGAAACGGTCCAAAGCCAGGAGGAGGAGCTTACCTGTTGAGT 680  
 Db 434 CAGCAAGAAAAGAAACGGTCCAAAGCCAGGAGGAGGAGCTTACCTGTTGAGT 375  
 QY 681 TTGCAATAGATCTTAATGAACCTACATCTGCTTTATGCAACCAAGTCTCTATGGGAGA 740  
 Db 374 TTGCAATAGATCTTAATGAACCTACATCTGCTTTATGCAACCAAGTCTCTATGGGAGA 315  
 QY 741 TGATAGATGTGACAAATGAACAGTGTCCAAATGAATGGTTTCACTTTTCATGTGTTTCAAC 800  
 Db 314 TGATAGATGTGACAAATGAACAGTGTCCAAATGAATGGTTTCACTTTTCATGTGTTTCAAC 255  
 QY 801 TTACCTATAAACCAAGGGGAAATGGTATTTGCCCAAGTGCAGGGGAGAGATAATGAGAAA 860  
 Db 254 TTACCTATAAACCAAGGGGAAATGGTATTTGCCCAAGTGCAGGGGAGAGATAATGAGAAA 195  
 QY 861 CAATGACAAAAGTACTGAAAAGACAAAAGAGATAGAAGATCGAGGTAGTAAAGCCAT 920  
 Db 194 CAATGACAAAAGTACTGAAAAGACAAAAGAGATAGAAGATCGAGGTAGTAAAGCCAT 135

QY 921 CCACATTTTAAAGGGTATTGCTCTTTATATATATGCTTTGCTTTTTCAGAAAATGTTTAA 980  
 Db 134 CCACATTTTAAAGGGTATTGCTCTTTATATATATGCTTTGCTTTTTCAGAAAATGTTTAA 75  
 QY 981 GGGTAATGCATAGACTATGCAATATATTTTAAATAGTATTAATAGTATGATGATTAATAA 1040  
 Db 74 GGGTAATGCATAGACTATGCAATATATTTTAAATAGTATTAATAGTATGATGATTAATAA 15  
 QY 1041 GTTGTGTTGACTTTG 1054  
 Db 14 GTTGTGTTGACTTTG 1

## RESULT 5

BE796780 793 bp mRNA linear EST 20-SEP-2000  
 DEFINITION 601587557F1 NIH\_MGC\_7 Homo sapiens cDNA clone IMAGE:3941655 5',  
 mRNA sequence.  
 ACCESSION BE796780  
 VERSION BE796780.1 GI:10218080  
 KEYWORDS EST.  
 SOURCE human.

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 793)

## AUTHORS

NIH-MGC http://mgc.nci.nih.gov/.

## TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

## JOURNAL

Unpublished (1999)

## COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: DCTD/DTF

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LEML at: image.llnl.gov

Plate: LLCM794 row: a column: 16

High quality sequence stop: 769.

## FEATURES

source

1. 793  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3941655"  
 /clone\_lib="NIH\_MGC\_7"  
 /tissue\_type="small cell carcinoma"  
 /cell\_line="MGC3"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: lung; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGCACGAG(G). Size-selected >500bp for average.  
 insert size 1.8kb. Library constructed by Ling Hong in  
 the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)."  
 BASE COUNT 240 a 177 c 230 g 146 t

## ORIGIN

Query Match 65.8%; Score 711; DB 12; Length 793;  
 Best Local Similarity 98.9%; Pred. No. 1e-98;  
 Matches 747; Conservative 0; Mismatches 5; Indels 3; Gaps 3;  
 QY 1 GCGGCGCGCGCGCGTGCATGTGCGGCTGCTGGATGCGGAGCGCGCGCGCGCGAT 60  
 Db 41 GCGGCGCGCGCGCGTGCATGTGCGGCTGCTGGATGCGGAGCGCGCGCGCGCGAT 100  
 QY 61 CGCGCAGATGTTAGGCGAGCAGCAGCAGCAACTGTACTCTCGGCTCGCTCTCTGACCGG 120  
 Db 101 CGCGCAGATGTTAGGCGAGCAGCAGCAGCAACTGTACTCTCGGCTCGCTCTCTGACCGG 160  
 QY 121 GGAGCGGAGCGCGCTGCTACCTGCTACGTGCAGGACTACCTTGTAGTGGGTGCTGCT 180

Db	161	GGAGCGGAGCGCGTCTCACCTTGGTACGAGTACGAGTACCTTGGTGGAGTCCGCT	222	
Qy	181	GCCCGGACATGCAGAGGACGCTGTCTGTCTGCGAGAGCTGGACAAACAATATCAAGA	240	
Db	221	GCCCGGACATGCAGAGGACGCTGTCTGTCTGCGAGAGCTGGACAAACAATATCAAGA	280	
Qy	241	AACGTTAAAGGAAATGATGATGCTTACGAAAAATATTAAGAAAGAAGATGATTTAAACCA	300	
Db	281	AACGTTAAAGGAAATGATGATGCTTACGAAAAATATTAAGAAAGAAGATGATTTAAACCA	340	
Qy	301	GAGAAACGCTACAGCAGCTTCTCCAGAGAGCAGCTAATTAATAGTCAAGTAATGGGAGA	360	
Db	341	GAGAAACGCTACAGCAGCTTCTCCAGAGAGCAGCTAATTAATAGTCAAGTAATGGGAGA	400	
Qy	361	TGAAAAAATACAGATTTGTTACACAAATGCTCCAAATTTGGTGGAAATCGGCAAGACAAT	420	
Db	401	TGAAAAAATACAGATTTGTTACACAAATGCTCCAAATTTGGTGGAAATCGGCAAGACAAT	460	
Qy	421	GGAGTTACACTACAGTGTGTTTCCAGATGCTGCTGAAAGTGAACGAGCCTCAGATAAAGC	480	
Db	461	GGAGTTACACTACAGTGTGTTTCCAGATGCTGCTGAAAGTGAACGAGCCTCAGATAAAGC	520	
Qy	481	AAAGATGGATTCAGGCCAACACAGAAAGATCTTCAAGAAAGACCCCGCAGCGGACGACAG	540	
Db	521	AAAGATGGATTCAGGCCAACACAGAAAGATCTTCAAGAAAGACCCCGCAGCGGACGACAG	580	
Qy	541	TGAAGCGCGTGATTTA-TGTCACATGCGAAATGGATTTGAAGACTGTGATGATCAGCCAC	599	
Db	581	TGAAGCGCGTGATTTA-TGTCACATGCGAAATGGATTTGAAGACTGTGATGATCAGCCAC	640	
Qy	600	CTAAGAAAAAGAAATCCAAGTCAGCAAGAAAAAGAAACGCTCCAAGGCCAAGCAGGAAA	659	
Db	641	CTAAGAAAAAGAAATCCAAGTCAGCAAGAAAAAGAAACGCTCCAAGGCCAAGCAGGAAA	700	
Qy	660	GGGAGCTTACCTGTTGAGTTTGCATAGATGCTTAATGAACCTACATAGCTGCTTATGCA	719	
Db	701	GGGAGCTTACCTGTTGAG-TTGCAATAGATGCTTAATGAACCTACATAGCTGCTTATGCA	759	
Qy	720	ACCAAGTGCTTATGGGAGATGATAGATGTGAC	754	
Db	760	ACCAAGTGCTT-GGGGAGATGATGCTGTC	793	
RESULT 6	BI548536	778 bp	linear	EST 05-SEP-2000
LOCUS	60319125F1 NIH_MGC_95	Homo sapiens	cdna clone	IMAGE:5262540 5',
DEFINITION	mRNA sequence.			
ACCESSION	BI548536			
VERSION	BI548536.1			
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	1 (bases 1 to 778)			
TITLE	NIH-MGC http://mgi.nci.nih.gov/.			
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)			
COMMENT	Unpublished (1999)			
	Contact: Robert Strausberg, Ph.D.			
	Email: cgapbs-r@mail.nih.gov			
	Tissue Procurement: Miklos Palkovits, M.D., Ph.D.			
	CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki			
	Toshiyuki and Piero Carninci (RIKEN)			
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
	DNA Sequencing by: Incyte Genomics, Inc.			
	Clone distribution: MGC clone distribution information can be			
	found through the I.M.A.G.E. Consortium/LLNL at:			
	http://image.llnl.gov			
	Plate: LLAM1661 row: n column: 13			
	High quality sequence stop: 775.			
	Location/Qualifiers			
	1..778			
FEATURES	source			



optic nerve, retina, Retina Foveal and Macular, RPE and Choroid"

/dev\_stage="fetal and adult"

/lab\_host="PH10B (Life Technologies) (T1 phage resistant)"

/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site.1: Ecor I; Site.2: Not I; UI-E-EJ1 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an Ecor I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are: fetal eyes, AGATCAAGA; lens, CGATTAGCGA; eye anterior segment, AATGCCGCAT; optic nerve, CCATTAAAGTG; retina, CCGCG; Retina Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).

TAG\_LIB=UI-E-EJ1

TAG\_TISSUE=Foveal and Macular Retina

TAG\_SEQ=GTCC"

BASE COUNT 164 a 148 c 110 g 253 t

ORIGIN

Query Match 62.2%; Score 671.8; DB 14; Length 675;

Best Local Similarity 99.7%; Pred. No. 9.6e-93;

Matches 673; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 398 GTGGAATTCGGGCAAGACAAATGGAGTTTACACTCACAGTGTTCACAGATCCTGCTGAA 457

DB 675 GTGGAATTCGGGCAAGACAAATGGAGTTTACACTCACAGTGTTCACAGATCCTGCTGAA 616

QY 458 AGTGAACGAGCTCAGATTAACCAAGATGGATTCCAGCCACCAAGAAAGATCTTCAAGA 517

DB 615 AGTGAACGAGCTCAGATTAACCAAGATGGATTCCAGCCACCAAGAAAGATCTTCAAGA 556

QY 518 AGACCCCGCAGCAGCGGACCAAGTGAAGCCGTGATTATGTACATGCAATGGGATT 577

DB 555 AGACCCCGCAGCAGCGGACCAAGTGAAGCCGTGATTATGTACATGCAATGGGATT 496

QY 578 GAAGACTGTGATGATCAGCCACTTAAGAAAAGAAATCAAGTCAGCAAGAAAAGAA 637

DB 495 GAAGACTGTGATGATCAGCCACTTAAGAAAAGAAATCAAGTCAGCAAGAAAAGAA 436

QY 638 CGCTCAAGGCCAAGCAGGAAGGAGCTTCACTGTGTGAGTTTGCATATAGTCTTAT 697

DB 435 CGCTCAAGGCCAAGCAGGAAGGAGCTTCACTGTGTGAGTTTGCATATAGTCTTAT 376

QY 698 GAACCTACATCTGCTTATGCAACCAAGTGTCTTATGGGAGATGATAGGATGTGACAT 757

DB 375 GAACCTACATCTGCTTATGCAACCAAGTGTCTTATGGGAGATGATAGGATGTGACAT 316

QY 758 GAACAGTGTCCAAATGAATGTTTCACTTTTCATGTGTTTCACTTACCTATTAACCAAG 817

DB 315 GAACAGTGTCCAAATGAATGTTTCACTTTTCATGTGTTTCACTTACCTATTAACCAAG 256

QY 818 GGGAAATGGTATTCGCCAAGTGCAGGGGAGATATAGAAAACAAATGGACAAAGTACT 877

DB 255 GGGAAATGGTATTCGCCAAGTGCAGGGGAGATATAGAAAACAAATGGACAAAGTACT 196

QY 878 GAAAAGACAAAAGATAGAGATCGAGGTAGTAAAGGCCATCCACATTTTAAAGGGTT 937

DB 195 GAAAAGACAAAAGATAGAGATCGAGGTAGTAAAGGCCATCCACATTTTAAAGGGTT 136

QY 938 ATTGTGCTTTTATATTCGTTTTCGTTTCAGAAAATGTTTAAAGGTAATGCAATAGAC 997

DB 135 ATTGTGCTTTTATATTCGTTTTCGTTTCAGAAAATGTTTAAAGGTAATGCAATAGAC 76

QY 998 TATGCAATAATTTTAAATCATTAGTATTAAATGCTATTAAAGTGTGTGTTGTTGAA 1057

|||||

Db 75 TATGCAATAATTTTAAATCATTAGTATTAAATGCTATTAAAGTGTGTGTTGTTGAA 16

QY 1058 AAAAAAATAATTTTAAATCATTAGTATTAAATGCTATTAAAGTGTGTGTTGTTGAA 1072

Db 15 AAAAAAATAATTTTAAATCATTAGTATTAAATGCTATTAAAGTGTGTGTTGTTGAA 1

RESULT 9

LOCUS BM719460

DEFINITION UI-E-EJ1-af-g-22-0-UI-r1 UI-E-EJ1 Homo sapiens cDNA clone

ACCESSION BM719460

VERSION BM719460.1 GI:19038113

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 664)

AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 9704477

COMMENT Contact: Soares, MB

Program for Rat Gene Discovery and Mapping

University of Iowa

451 Eckstein Medical Research Building Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: msoares@blue.weeg.uiowa.edu

Tissue Procurement: Dr. Gregg Hageman

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).

Seq primer: M13 Reverse.

FEATURES

Location/Qualifiers

1..664

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="UI-E-EJ1-af-g-22-0-UI"

/clone\_lib="UI-E-EJ1"

/tissue\_type="fetal eyes, lens, eye anterior segment, optic nerve, retina, Retina Foveal and Macular, RPE and Choroid"

/dev\_stage="fetal and adult"

/lab\_host="PH10B (Life Technologies) (T1 phage resistant)"

/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site.1: Ecor I; Site.2: Not I; UI-E-EJ1 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an Ecor I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are: fetal eyes, AGATCAAGA; lens, CGATTAGCGA; eye anterior segment, AATGCCGCAT; optic nerve, CCATTAAAGTG; retina, CCGCG; Retina Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

BASE COUNT 241 a 111 c 149 g 162 t

ORIGIN

Query Match 61.4%; Score 663; DB 14; Length 664;

Best Local Similarity 99.8%; Pred. No. 2.1e-91;

Matches 663; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 377 GTTACAAATGCTCGAATGGTGGAAATCGGCAAGACAAATGGAGTTACACTCACAG 436
|
|
|
Db 1 GTTACAAATGCTCGAATGGTGGAAATCGGCAAGACAAATGGAGTTACACTCACAG 60
|
|
|
QY 437 TGTTCACAGATCCTGCTGAAAGTGAACAGACCTCAGATAAGCAAGATGGATTCACGC 496
|
|
|
Db 61 TGTTCACAGATCCTGCTGAAAGTGAACAGACCTCAGATAAGCAAGATGGATTCACGC 120
|
|
|
QY 497 CAACCAAGAGATCTTCAAGAACCCCGCAGCAGCGGACCACTGAAAGCCGTGATTAA 556
|
|
|
Db 121 CACCCAGAAAGATCTTCAAGAACCCCGCAGCAGCGGACCACTGAAAGCCGTGATTAA 180
|
|
|
QY 557 TGTACATGCAATGGATTTGAAGACTGTGATGATCAGCCACCTAAAGAAAGAAATCC 616
|
|
|
Db 181 TGTACATGCAATGGATTTGAAGACTGTGATGATCAGCCACCTAAAGAAAGAAATCC 240
|
|
|
QY 617 AAGTCAGCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAATCC 676
|
|
|
Db 241 AAGTCAGCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAATCC 300
|
|
|
QY 677 GAGTTTCCATAGATCCTTAATGAACCTTACATCTGCTTATGCAACCAAGTCTTATGGG 736
|
|
|
Db 301 GAGTTTCCATAGATCCTTAATGAACCTTACATCTGCTTATGCAACCAAGTCTTATGGG 360
|
|
|
QY 737 GAGATGATAGATGTGACAAATGAACAGTGTCCAAATGGAATGGTTTTCATGATGTT 796
|
|
|
Db 361 GAGATGATAGATGTGACAAATGAACAGTGTCCAAATGGAATGGTTTTCATGATGTT 420
|
|
|
QY 797 TCACCTACCTATTAACCAAGGGGAATGGTATTTGCCCAAGTGCAGCGGGAGATAATGAG 856
|
|
|
Db 421 TCACCTACCTATTAACCAAGGGGAATGGTATTTGCCCAAGTGCAGCGGGAGATAATGAG 480
|
|
|
QY 857 AAAACAATGACAAAAGTACTGAAAGACAAAAGGATAGAGATCGAGTAGTAAGG 916
|
|
|
Db 481 AAAACAATGACAAAAGTACTGAAAGACAAAAGGATAGAGATCGAGTAGTAAGG 540
|
|
|
QY 917 CCATCCACATTTTAAAGGGTATTTGCTTTTATATATTCGTTTTCAGAAAATGT 976
|
|
|
Db 541 CCATCCACATTTTAAAGGGTATTTGCTTTTATATATTCGTTTTCAGAAAATGT 600
|
|
|
QY 977 TTTAGGTAATGCAATAGACTATGCAATATTTTATCATTTAGTATTAATGGTGATT 1036
|
|
|
Db 601 TTTAGGTAATGCAATAGACTATGCAATATTTTATCATTTAGTATTAATGGTGATT 660
|
|
|
QY 1037 AAAA 1040
|
|
|
Db 661 AAAA 664
```

```
RESULT 10
BI495873/c
LOCUS
DEFINITION df121d04.v1 Morton Fetal Cochlea Homo sapiens cDNA clone
IMAGE:2540190.3', mRNA sequence.
ACCESSION BI495873
VERSION BI495873.1 GI:15335217
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 653)
Robertson, N.G., Khetarpal, U., Gutierrez-Espeleta, G.A., Bieber, F.R.,
and Morton, C.C.
Isolation of novel and known genes from a human fetal cochlear cDNA
library using subtractive hybridization and differential screening
Genomics 23, 42-50 (1994)
95130111
MEDLINE
COMMENT Contact: Morton, C. C.
Departments of Pathology and Obstetrics, Gynecology and
Reproductive Biology
Brigham and Women's Hospital
```

75 Francis Street, Harvard Medical School, Boston, MA 02115, USA  
Tel: 617 732 7980  
Fax: 617 732 6996  
Email: cmorton@rics.bwh.harvard.edu  
DNA sequencing and analyses were performed by National Institutes  
of Health Intramural Sequencing Center (NISC; see  
<http://www.nisc.nih.gov>).  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
Plate: L1AM6327 row: G column: 7  
Seq primer: T7 primer.  
Location/Qualifiers  
1. 653

FEATURES  
source

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2540190"  
/clone\_lib="Morton Fetal Cochlea"  
/tissue\_type="cochlea"  
/dev\_stage="16-22 week fetus"  
/lab\_host="SOLR cells (kanamycin resistant)"  
/note="Organ: ear; Vector: pBluescript SK-; Site: 1; EORI;  
Site 2: XhoI; Reference: Genomics 23, 42-50 (1994) Cloned  
unidirectionally. Primer: Oligo df. Fetal cochlea, normal.  
37% of inserts <0.5 kb, 56% 0.5-1.0 kb, 7% >1 kb. Uni-ZAP  
XR Vector. Library constructed by N. Robertson, C. Morton.  
-5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor  
sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3' "

BASE COUNT 160 a 146 c 106 g 241 t  
ORIGIN

Query Match 58.9%; Score 636.4; DB 13; Length 653;  
Best Local Similarity 99.8%; Pred. No. 2.3e-87;  
Matches 637; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 435 AGTGTTCACAGATCCTGCTGAAAAGTGAACAGAGCTCAGATAAGCAAGATGGATTCGA 494  
|  
|  
|  
Db 653 AGTGTTCACAGATCCTGCTGAAAAGTGAACAGAGCTCAGATAAGCAAGATGGATTCGA 594  
|  
|  
|  
QY 495 GCCAACCAAGAAATGCTTCAAGAAAGACCCCGCAGCAGCGGACCACTGAAAGCGGTGATT 554  
|  
|  
|  
Db 593 GCCAACCAAGAAATGCTTCAAGAAAGACCCCGCAGCAGCGGACCACTGAAAGCGGTGATT 534  
|  
|  
|  
QY 555 TATGTCATGCGCAATGGGATTTGAAGACTGTGATGATCAGCCACCTAAAGAAAGAAAT 614  
|  
|  
|  
Db 533 TATGTCATGCGCAATGGGATTTGAAGACTGTGATGATCAGCCACCTAAAGAAAGAAAT 474  
|  
|  
|  
QY 615 CCAAGTCAGCAAGAAAGAAAGAAAGCTCCAGGCGGACCAAGGAAAGGAAAGCTTCACCTG 674  
|  
|  
|  
Db 473 CCAAGTCAGCAAGAAAGAAAGAAAGCTCCAGGCGGACCAAGGAAAGGAAAGCTTCACCTG 414  
|  
|  
|  
QY 675 TTGAGTTTGCATAGATCCTTAATGAACCTACATCTGCTTATGCAACCAAGTGTCTTATG 734  
|  
|  
|  
Db 413 TTGAGTTTGCATAGATCCTTAATGAACCTACATCTGCTTATGCAACCAAGTGTCTTATG 354  
|  
|  
|  
QY 735 GGGAGATGATAGGATGTGACAAATGAACAGTGTCCAATTTGAATGTTTTCACATTTTCATGTG 794  
|  
|  
|  
Db 353 GGGAGATGATAGGATGTGACAAATGAACAGTGTCCAATTTGAATGTTTTCACATTTTCATGTG 294  
|  
|  
|  
QY 795 TTTTCATTTACCTATAACCAAGGGAAATGGTATTTGCCCAAGGAGGAGGAGAGTAATG 854  
|  
|  
|  
Db 293 TTTTCATTTACCTATAACCAAGGGAAATGGTATTTGCCCAAGGAGGAGGAGAGTAATG 234  
|  
|  
|  
QY 855 AGAAACAATGGACAAAGTACTGAAAGACAAAGGATAGAGATCGAGGAGTAGTAAA 914  
|  
|  
|  
Db 233 AGAAACAATGGACAAAGTACTGAAAGACAAAGGATAGAGATCGAGGAGTAGTAAA 174  
|  
|  
|  
QY 915 GGCCATCCACATTTTAAAGGGTATTTGCTCTTTTATATAATTCGTTTTCAGAAAAT 974  
|  
|  
|  
Db 173 GGCCATCCACATTTTAAAGGGTATTTGCTCTTTTATATAATTCGTTTTCAGAAAAT 114  
|  
|  
|  
QY 975 GTTTTAGGGTAAATGCATAGACTATGCAATTAATTTTAAATCATTTAGTATTAATGGTGA 1034  
|  
|  
|  
Db 113 GTTTTAGGGTAAATGCATAGACTATGCAATTAATTTTAAATCATTTAGTATTAATGGTGA 54



**BASE COUNT**





<b>DEFINITION</b>						
ba59f06.x1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:2900867 3' similar to TR:O95698 O95698 INGIILP.; mRNA sequence.						
<b>ACCESSION</b>						
AW674008						
<b>VERSION</b>						
AW674008.1 GI:7539243						
<b>KEYWORDS</b>						
EST.						
<b>SOURCE</b>						
human.						
<b>ORGANISM</b>						
Homo sapiens						
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
<b>REFERENCE</b>						
1 (bases 1 to 574)						
NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .						
<b>AUTHORS</b>						
National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)						
<b>TITLE</b>						
Other_ESTs: ba59f06.y1						
<b>JOURNAL</b>						
Contact: Robert Strausberg, Ph.D.						
<b>COMMENT</b>						
Email: cgapbs-re@mail.nih.gov						
Tissue Procurement: ATCC						
CDNA Library Preparation: Life Technologies, Inc.						
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)						
DNA Sequencing by: Washington University Genome Sequencing Center						
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:						
<a href="http://image.llnl.gov/image/html/iresources.shtml">image.llnl.gov/image/html/iresources.shtml</a>						
Seq primer: -40UP from Gibco						
High quality sequence stop: 420.						
<b>FEATURES</b>						
Location/Qualifiers						
1..574						
<code>/organism="Homo sapiens"</code>						
<code>/db_xref="taxon:9606"</code>						
<code>/clone="IMAGE:2900867"</code>						
<code>/cell_line="NIH_MGC_10"</code>						
<code>/cell_lines="MGC36"</code>						
<code>/lab_host="DH10B"</code>						
<code>/note="Organ: cervix; Vector: pCMV-SPORT6; Site.1: NotI; Site.2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."</code>						
<b>BASE COUNT</b>		147 a	125 c	90 g	211 t	1 others
<b>ORIGIN</b>						
Query Match               52.2%; Score 563.4; DB 10; Length 574;						
Best Local Similarity 98.8%; Pred. No. 2.8e-76;						
Matches 567; Conservative 0; Mismatches 7; Indels 0; Gaps 0						
QY	488	GATTCAGCACAACCAGAAGATCTTCAAGAGACCCCGCAGGACGGGACCATGAAGC	547			
DB	574	GATTCAGCCACACAGANAGATTTCAAAGAGACCCCGCAGGACGGGCCACTTAAGC	515			
QY	548	CGTGATTATTGTACATGGCAAATGGGATTGAAGCTGTGATGATCAGCCACCTAAAGAA	607			
DB	514	CGTGATTATTGTACATGGCAAATGGGATTGAAGCTGTGATGATCAGCCACCTAAAGAA	455			
QY	608	AGAAATCCAAGTCAGCAAAAAGAAAGAACCGTTCCAAAGCCCAAGCAGGAAGGAAGCT	667			
DB	454	AGAAATCCAAGTCAGCAAAAAGAAAGAACCGTTCCAAAGCCCAAGCAGGAAGGAAGCT	395			
QY	668	TCACCTGTGTAGTTTGCATAATAGATCCCTAATGAACCTACATCTCTTATGCAACCAAGTG	727			
DB	394	TCACCTGTGTAGTTTGCATAATAGATCCCTAATGAACCTACATCTCTTATGCAACCAAGTG	335			
QY	728	TCTTATGGGGAGATGATAGGATGTGACAATGAACAGTGTCCAATTGAATGGTTTCACCTT	787			
DB	334	TCTTATGGGGAGATGATAGGATGTGACAATGAACAGTGTCCAATTGAATGGTTTCACCTT	275			
QY	788	TCATGTGTTTCACTTACCTATAAACCAAGGGGAATGGTATTGCCCAAGTGCAGGGGA	847			
DB	274	TCATGTGTTTCACTTACCTATAAACCAAGGGGAATGGTATTGCCCAAGTGCAGGGGA	215			
QY	848	GATAATGAGAAAAACAATGCACAAAAGTACTGAAAAGACAAAAAAGATAGAGATCGAG	907			
DB	214	GATATGAGAAAAACAATGCACAAAAGTACTGAAAAGACAAAAAAGATAGAGATCGAG	155			
QY	908	TAGTAAGGCCATCCACATTTTTAAAGGGTATTTGTCTTTTATATATATTCGTTTGTCTTC	967			

Db	154	TAGTAAAGGCCATCCACATTTTAAAGGGTATTGTGCTTTTATATAAATTCGTTGCTTC	95
Qy	968	AGAAAATGTTTTAGGCGTAATGCATAAGACTATGCAATAATTTTAAATCATAGTATTAA	1027
Db	94	AGAAAATGTTTTAGGCGTAATGCATAAGACTATGCAATAATTTTAAATCATAGTATTAA	35
Qy	1028	TGCTGTATTAAAAAGTTTGTGTACTTTGAAAAAA	1061
Db	34	TGCTGTATTAAAAAGTTTGTGTACTTTGAAAAAA	1
RESULT 15			
BE839427/c			
LOCUS	BE839427	564 bp	linear EST 22-SEP-2000
DEFINITION	RC3-FN0143-190700-022-d12 FN0143	Homo sapiens	cdna, mRNA sequence.
ACCESSION	BE839427		
VERSION	BE839427.1	GI:10271805	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
	1 (bases 1 to 564)		
	Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,		
	Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,		
	Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,		
	Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare		
	, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and		
	Simpson, A.J.		
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed		
	sequence tags		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496		(2000)
MEDLINE	20202663		
COMMENT	Contact: Simpson A.J.G.		
	Laboratory of Cancer Genetics		
	Ludwig Institute for Cancer Research		
	Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,		
	Brazil		
	Tel: +55-11-2704922		
	Fax: +55-11-2707001		
	Email: asimpson@ludwig.org.br		
	This sequence was derived from the FAPESP/LICR Human Cancer Genome		
	Project. This entry can be seen in the following URL		
	( <a href="http://www.ludwig.org.br/scripts/gethtml2.pl?tl1=&amp;tl2=RC3-FN0143-190700-022-d12&amp;tl3=2000-07-19&amp;tl4=1">http://www.ludwig.org.br/scripts/gethtml2.pl?tl1=&amp;tl2=RC3-FN0143-190700-022-d12&amp;tl3=2000-07-19&amp;tl4=1</a> )		
	Seq primer: puc 18 forward		
	High quality sequence start: 18		
	High quality sequence stop: 564.		
FEATURES	Location/Qualifiers		
source	1. 564		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone_lib="FN0143"		
	/dev_stage="Adult"		
	/note="Organ: prostate_normal; Vector: puc18; Site:1: SmaI		
	products derived from ORESTES PCR (U.S. Letters Patent		
	application No. 196,716 - Ludwig Institute for Cancer		
	Research) profiles into the puc 18 vector. Reverse		
	transcription of tissue mRNA and cDNA amplification were		
	performed under low stringency conditions."		
BASE COUNT	108 a 156 c 121 g 179 t		
ORIGIN			
	51.3%; Score 554.2; DB 12; Length 564;		
Query Match	99.5%; Pred. No. 7e-75;		
Best Local Similarity	99.5%;		
Matches	556; Conservative	0; Mismatches	3; Indels
		0; Gaps	0;
Qy	38	GAGCGCGCGCAGCGCGGATTCGCGAGATTTAGGCGAGCAGCAGCAGCACTGTAC	97
Db	564	GAGCGCGCGCAGCGCGGATTCGCGAGATTTAGGCGAGCAGCAGCAGCACTGTAC	505

Qy	98	TCGTCGGCTCGCTCCTGACCGGGAGCGGAGCGGGTGTCTACCTGCTAGCTGCAGGAC	157
Db	504	TCGTCGGCTCGCTCCTGACCGGGAGCGGAGCGGGTGTCTACCTGCTAGCTGCAGGAC	445
Qy	158	TACCTTGAGTCGGTGGAGTCGCTGCCACACATGACAGGAGCGTCTGTGCTGGGA	217
Db	444	TACCTTGAGTCGGTGGAGTCGCTGCCACACATGACAGGAGCGTCTGTGCTGGGA	385
Qy	218	GAGCTGGACACAAATATCAAGAAACGTTAAAGGAAATTTGATGATGTCTAGGAAAAATAT	277
Db	384	GAGCTGGACACAAATATCAAGAAACGTTAAAGGAAATTTGATGATGTCTAGGAAAAATAT	325
Qy	278	AAGAAAGAGATGATTTAAACCAAGAAACGTTACAGCAGCTTCTCCAGAGAGCACTA	337
Db	324	AAGAAAGAGATGATTTAAACCAAGAAACGTTACAGCAGCTTCTCCAGAGAGCACTA	265
Qy	338	ATTAATAGTCAAGAAATGGGAGATGAAATAATACAGATTGTTACACAAATGCTCGAATTG	397
Db	264	ATTAATAGTCAAGAAATGGGAGATGAAATAATACAGATTGTTACACAAATGCTCGAATTG	205
Qy	398	GTGAAAAATCGGGCAAGACAAATGGAGTTACACTCACAGTGTTCCTCAAGATCCTGCTGAA	457
Db	204	GTGAAAAATCGGGCAAGACAAATGGAGTTACACTCACAGTGTTCCTCAAGATCCTGCTGAA	145
Qy	458	AGTGAACGAGCCTCAGATAAAGCAAGATGGATTCCAGCCCAACCAAGAAAGATCTTCAAGA	517
Db	144	AGTGAACGAGCCTCAGATAAAGCAAGATGGATTCCAGCCCAACCAAGAAAGATCTTCAAGA	85
Qy	518	AGACCCCGCAGCGGAGCGGACCGAGTGAAGCGGTGATTATGTCATGCGCAAAATGGGATT	577
Db	84	AGACCCCGCAGCGGAGCGGACCGAGTGAAGCGGTGATTATGTCATGCGCAAAATGGGATT	25
Qy	578	GAAGACTGTGATGATCAGC	596
Db	24	GAAGACTGTGATGAGCACC	6

Search completed: June 20, 2003, 02:40:38  
Job time : 1538 secs



Result No.	Score	Query Match	Length	DB	ID	Description
1	1080	100.0	1080	22	AA012783	Tumour suppressor
2	1046	96.9	1078	22	AA0208596	Human INGL1 encodi
3	838.4	77.6	840	20	AA008595	Human INGL1 encodi
4	825.2	76.4	1153	21	AA247473	Human tumour suppressor
5	332.2	30.8	346	21	AA028459	Human secreted protein
6	325	30.1	325	21	AA247474	Human tumour suppressor
7	253.6	23.5	2817	21	AA053790	Murine p37ING1 codi
8	237.4	22.0	911	21	AA053792	Human p37ING1 codi
9	236	21.9	1533	22	AA028478	Nucleotide sequence

XXI



\_\_\_\_\_





Db 720 AAAGAAACGCTCCAGGCCAAGCAGGAAAGGAGCTTACCTGTTGAGTTGCAATAGA 779  
 Qy 691 TCCTAATGAACCTACATACCTGCTTATGCAACCAAGTCTTATGGGAGATGATAGGATG 750  
 Db 780 TCCTAATGAACCTACATACCTGCTTATGCAACCAAGTCTTATGGGAGATGATAGGATG 839  
 Qy 751 TGCAATGACAGTGTCCATTCGAATGATGTTTCACTTTTCACTTTTCACTTTTCACTTTA 810  
 Db 840 TGCAATGACAGTGTCCATTCGAATGATGTTTCACTTTTCACTTTTCACTTTTCACTTTA 899  
 Qy 811 ACCAAGGGGAAATGGTATTGCCCAAGTGCAGGGAGATTAATGAGAAACAAATGGACAA 870  
 Db 900 ACCAAGGGGAAATGGTATTGCCCAAGTGCAGGGAGATTAATGAGAAACAAATGGACAA 959  
 Qy 871 AAGTACTGAAAGACAAAGGATGAGATGATGATGATGATGATGATGATGATGATGATGAT 930  
 Db 960 AAGTACTGAAAGACAAAGGATGAGATGATGATGATGATGATGATGATGATGATGATGAT 1019  
 Qy 931 AAGGTTTATTTGCTTTTATATATTCGTTTTCGTTTTCGTTTTCGTTTTCGTTTTCGTTT 990  
 Db 1020 AAGGTTTATTTGCTTTTATATATTCGTTTTCGTTTTCGTTTTCGTTTTCGTTTTCGTTT 1079  
 Qy 991 ATAAGACTATGCAATATTTTATATATTCGTTTTCGTTTTCGTTTTCGTTTTCGTTTTCGTTT 1050  
 Db 1080 ATAAGACTATGCAATATTTTATATATTCGTTTTCGTTTTCGTTTTCGTTTTCGTTTTCGTTT 1139  
 Qy 1051 TTTGAAAAA 1064  
 Db 1140 TTTGAAAAA 1153

RESULT 5

AAAC28459  
 ID AAC28459 standard; cDNA; 346 BP.  
 AC AAC28459;  
 DT 06-OCT-2000 (first entry)  
 DE Human secreted protein 5' EST, SEQ ID NO: 32534.  
 DE Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 KW gene therapy; chromosome mapping; ss.  
 KW Homo sapiens.  
 OS EP1033401-A2.  
 PN 06-SEP-2000.  
 PD 21-FEB-2000; 2000EP-0200610.  
 PF 26-FEB-1999; 99US-0122487.  
 PR (GEST ) GENSET.  
 PA Dumas Milne Edwards J, Duclert A, Giordano J;  
 PI WPI; 2000-500381/45.  
 DR New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
 XX diagnostic, forensic, gene therapy and chromosome mapping procedures -  
 XX Claim 1; SEQ ID 32534; 71pp + CD-ROM; English.  
 XX The present sequence is one of a large number of 5' ESTs derived from  
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively  
 CC identified within the present sequence. The 5' ESTs were prepared from  
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST  
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)  
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA  
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences

CC derived from the 5' ends of mRNAs and even in those cases where longer  
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.  
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be  
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used  
 CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.  
 CC They are used to obtain upstream regulatory sequences and to design  
 CC expression and secretion vectors.  
 XX  
 SQ Sequence 346 BP; 130 A; 63 C; 80 G; 70 T; 3 other;  
 Query Match 30.88; Score 332.2; DB 21; Length 346;  
 Best Local Similarity 99.18; Pred. No. 3.2e-66;  
 Matches 343; Conservative 1; Mismatches 1; Indels 1; Gaps 1;  
 Qy 254 ATTGATGATGCTACGAAAAATATAAGAAAGAGATGATTTAAACCAAGAAACGCTCTA 313  
 Db 1 ATTGATGATGCTACGAAAAATATAAGAAAGAGATGATTTAAACCAAGAAACGCTCTA 60  
 Qy 314 CAGCAGCTTCTCCAGAGAGCAGCTAAATTAATAGTCAAGATTTGGGAGATGAAAAATACAG 373  
 Db 61 CAGCAGCTTCTCCAGAGAGCAGCTAAATTAATAGTCAAGATTTGGGAGATGAAAAATACAG 120  
 Qy 374 ATTGTTACACAAATGCTCGAATTTGGTGAATTCGGCAAGACAAATGGAGTTACACTCA 433  
 Db 121 ATTGTTACACAAATGCTCGAATTTGGTGAATTCGGCAAGACAAATGGAGTTACACTCA 180  
 Qy 434 CAGTGTTCACCAAGATCCTGCTGAAAGTGAACGAGCCTCAGATAAAGCAAGATGGATTCC 493  
 Db 181 CAGTGTTCACCAAGATCCTGCTGAAAGTGAACGAGCCTCAGATAAAGCAAGATGGATTCC 240  
 Qy 494 AGCCAACCAAGAGATCTTCAAGAAAGACCCCGCAGCGGACCGAGTGAAGCCGTGAT 553  
 Db 241 AGCCAACCAAGAGATCTTCAAGAAAGACCCCGCAGCGGACCGAGTGAAGCCGTGAT 300  
 Qy 554 TT-ATGTCACATGGCAATGGGATTGAAGACTGTGATGATCAGCCA 598  
 Db 301 THNNTGTACATGGCAATGGGATTGAAGACTGTGATGATCAGCCA 346

RESULT 6

AAZ47474/C  
 ID AAZ47474 standard; cDNA; 325 BP.  
 XX  
 AC AAZ47474;  
 DT 14-MAR-2000 (first entry)  
 DE Human tumour suppressor (TUSUP) fragment nucleotide sequence.  
 DE Tumour suppressor protein; TUSUP; human; cancer; treat; prevent;  
 KW reproductive tract; gastrointestinal tract; immune system; ss.  
 KW Homo sapiens.  
 OS WO9961612-A1.  
 PN 02-DEC-1999.  
 PD 20-MAY-1999; 99WO-US11136.  
 PF 28-MAY-1998; 98US-0086359.  
 PR (INCY-) INCYTE PHARM INC.  
 PA Tang YT, Corley NC, Patterson C;  
 PI WPI; 2000-062711/05.  
 DR New human tumour suppressor protein for treating cancer, particularly of  
 XX reproductive and gastrointestinal tracts or immune system -  
 XX Disclosure; Page 64; 71pp; English.  
 XX



xx cc xx

RESULT 8  
AAA53792  
ID AAA53792 standard; CDNA; 911 BP.

XX	
AC	AAA53792;
XX	
DT	19-DEC-2000 (first entry)

XX DE Human P37ING1 coding sequence.

KW p53; tumour; cancer; detection; antibody; hybridisation; exon 1b;  
 KW ING1; ing1; p33ING1; p37ING1; oncogene; gene therapy; diagnosis;  
 KW proliferation disorder; transformation; transformed cell; human;

OS Homo sapiens.

PN WO200046370-A1.

PD 10-AUG-2000.

04-FEB-2000; 2000WO-US02959.

PR 04-FEB-1999; 99US-0118941.

PA (UNII ) UNIV ILLINOIS FOUND.

PI Gudkov A, Zeremski M, Gurova KV, Grigorian IA;

DR WPI; 2000-491278/43.

DR P-PSDB; AAY97244.

PT Detecting nucleic acid encoding exon 1b of *ing1*, useful for diagnosing  
PT and treating cancer, comprises contacting sample with isolated nucleic  
PT acid comprising sequence of exon 1b and detecting hybridized products

PS Disclosure; Page 122-123; 134pp; English.

Query Match	22.0%;	Score 237.4;	DB 21;	Length 911;
Best Local Similarity	59.0%;	Pred. No. 1.9e-44;		

Y 145 CTACGTGCAGGACTACCTTGAGTGGGTGGCTGCTGCCCCACGACATGCAGAGGAACGT 204

**b** 86 CTATGTGGAGGACTACCTGGACTCCATCGAGTCCCTGCCTTTCGACTTGCAGAGAAATGT 145

Y 205 GTCGTGCTGCGAGAGCTGGACAACAAATATCAAGAAACGTTAAAGGAAATTGATGATGT 264

b 146 CTCGCTGATCGGGGAGATCGACGCCGAAATACCAAGAGATCCTGAAGGAGCTAGACGAGTG 205

265 CTACGAAAAATATAAGAAAGAAGATGATTAAACCAGAGAAACGCTCTACAGCAGCTTCT 324

b  
206 CTACGAGCGCTTCAGTCGCGAGACAGACGGGGCGGATGCTGCACCTGTGT 265

325 CCAGAGAGCACTAATTATAGTCAAGAATTGGGAGATGAAAAATACAGATTGTTACACA 384

b 266 GCAGCGCGCTGATCCGCAGCCAGGAGCTGGGCGACGAGAAGATCCAGATCGTGAGCCA 325

385 AATGCTCGAATTGGTGGAAATCGGGCAAGACAAATGGAGTTACACTCACAGTGTTCCA 444

b . 326 GATGGTGAGCTGGTGAGAACCGCACGGCAGGTGGACAGCCACGTGGAGCTGTTCTGA 385

445 AGATC-----CTGCTGAAAGTGAACGAGCCTCAGATA 476

b 386 GGCGCAGGAGCTGGGCGACACAGTGGGCAACAGCGGAAGGTTGGCGGCACAGGCC 445

477 AAGCAAAGATGGATTCCAGCCAACCAGAAAGATCTTCAAGAAGACCCCGCAG--GCAGCG 534

b 446 CAAATGGCGATGCGGTAGCGCAGTCTGACAAAGCCCAACAGCAAGCGCTCACGGCGGACGG 505

535 GACCAGTGAAAGCCGTGATTATGTCACATGGCAAATGGGATTGAAGACTGTGATGATCA 594

b

506 CAACAACGAGAACCGTGAGAACGCGTCCAGCAACCACGACCACGACGCGCCTCGGG 565

595 GCCACCTAAAGAAAGAAATCCAAGTCAGCAAGAAAGAAACGGCTCCAGGCCAAGCA 654







Isolated cancer associated nucleic acid molecule identified by SEREX (serological identification of antigens by recombinant expression cloning) technique; useful in nucleic acid based therapies to treat cancer -

Example 4; Page 44; 62pp; English.

The present sequence encodes a human cancer associated antigen. The sequence is the wildtype of the INGI gene, which is a tumour suppressor gene candidate. The cancer associated antigen polynucleotides and polypeptides are useful for screening for the possible presence of a pathological condition in a subject such as cancer. The cancer associated antigen polypeptides are useful for producing vaccines.

Sequence 1143 BP; 289 A; 291 C; 366 G; 197 T; 0 other;

```

Query Match          17.0%; Score 183.8; DB 22; Length 1143;
Best Local Similarity 57.0%; Pred. No. 3.4e-32;
Matches 394; Conservative 0; Mismatches 267; Indels 30; Gaps 2;

QY 237 AAGAAAGCTTAAAGGAAATTTGATGCTACGAAATAATAAGAAAGAGATGATTTAA 296
DB 194 AACAGATCTCTGAAGAGCTAGACGAGTCTTACGAGCGCTTCACTCGCGAGACACGCGG 253
QY 297 ACCAGAGAACGCTCTACAGCAGCTTCTCCAGAGAGCACTAATTAATAGTCAAGATTGG 356
DB 254 CGCAGAGCGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 313
QY 357 GAGATGAAATAATACAGATCTTACACAAATGCTCGAATGCTGTAATAATCGGCAAGAC 416
DB 314 GCGAGAGAGATCCAGATCTGAGCGAGATGTTGGAGCTGTTGGAGTCCGACCGCAGCGGC 373
QY 417 AAATGGAGTTACATACAGTGTGTTTCCAAAGATC----- 449
DB 374 AGTGGACAGCCACGTGAGCTGTTTCGAGGCGCAGCAGGAGCTGGCGGACACAGTGGCA 433
QY 450 -CTGCTGAACTGAACAGCCTCAGATTAACAAAGATGGATTCAGCCCAACCAAGAAGA 508
DB 434 ACAGCGGCAAGTTGGCGGCGGACAGGCGCCATGCGATGCGGTAGCGCAGTCTGACAAAG 493
QY 509 TCTTCAAGAA--GACCCGCGAGCGGCGGACGAGCCAGTGAAGCCGTGATTTATGTACATGG 566
DB 494 CCAACAGCAGCGCTACGCGCGGCGGCGGACGAGAACCGGTGAGAACCGGTCCAGCA 553
QY 567 CAATGGGATTGAAGACTGTGATGATCAGCCACCTAAAGAAAGAAATCCAAAGTCAGCAA 626
DB 554 ACCACGACGACGAGCGGCGCTCGGCGCACACCCAGGAGAGAGGCGCAAGACCTCA 613
QY 627 AGAAAGAGACGCTCCAGGCGCAAGCAGGAGGAGGAGCTTACCTGTTGAGTTGGCA 686
DB 614 AGAAGAGAGACGCTCCAAAGGCGCAAGGCGGAGGAGGCGCTCCCTGCCACCTCCCA 673
QY 687 TAGATCTTAATGAACCTACATCTGTTATGCAACCAAGTCTTATGGGAGATGATAG 746
DB 674 TCGACCCCAACGACCCACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 733
QY 747 GATGTGCAATGAACAGTGTCCCAATGTAATGTTTCACTTTTCACTTTTCACTTTACT 806
DB 734 GCTGCGACACGACGAGTGCCCATCGAGTGTGTTCCACTTCTGCTGCTGCTGCTGCTGCT 793
QY 807 ATAAACCAAGGGGAATGTTATGTTCCCAAGTGCAGGGGAGATTAATGAGAAACAATGG 866
DB 794 ATAAACCAAGGGGAATGTTATGTTCCCAAGTGCAGGGGAGATTAATGAGAAACAATGG 853
QY 867 ACAAGGACTGAAAGACAAAGAGGATAG 897
DB 854 ACAAGGCTTGAGAAATCCAAAGAGAGAG 884

```

RESULT 13

AAT69651

ID AAT69651 standard; cDNA; 1902 BP.

XX

AC AAT69651;

XX 27-AUG-1997 (first entry)

DE Tumour suppressor gene INGI partial cDNA.

XX Tumour suppressor gene; INGI; p33INGI; breast cancer; brain cancer; diagnosis; gene therapy; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 109..741

FT /\*tag= a

PN W09721809-A1.

PD 19-JUN-1997.

XX 06-DEC-1996; 96WO-CA00819.

XX 15-NOV-1996; 96US-0751230.

XX 08-DEC-1995; 95US-0569721.

XX (UYTE-) UNIV TECHNOLOGIES INT INC.

XX Garkavstev I, Riabowol K;

XX WPI: 1997-332781/30.

XX P-FSDB; AAW19118.

XX Isolated tumour suppressor gene, INGI - useful to develop products for inhibiting or increasing cell proliferation, in particular for treatment or diagnosis of cancer

XX Claim 1; Fig 2; 63pp; English.

XX A partial cDNA clone (AAT69651), designated INGI, codes for a novel tumour suppressor protein p33INGI (AAW19118) that is a potent inhibitor of cell growth. It was isolated by subtractive

CC hybridisation between normal mammary and transformed epithelial

CC cDNAs, isolation of an antisense INGI cDNA insert that caused

CC increased cell proliferation, and use of the insert to screen

CC normal human fibroblast and HeLa cDNA libraries. A complete INGI

CC sequence (AAT69652) was obtained by 5'RACE. INGI is localised to the

CC 13q33-34 chromosome region, to which a number of human cancers have

CC been mapped. INGI nucleic acids can be used in the diagnosis of

CC breast cancer; a decreased level of INGI mRNA indicates cancerous

CC cells. They can also be used in gene therapy methods to block the

CC proliferation of cancer cells.

XX Sequence 1902 BP; 574 A; 391 C; 461 G; 476 T; 0 other;

Query Match 17.0%; Score 183.8; DB 18; Length 1902;

Best Local Similarity 57.0%; Pred. No. 4.1e-32;

Matches 394; Conservative 0; Mismatches 267; Indels 30; Gaps 2;

QY 237 AAGAAACCTTAAAGGAAATTTGATGCTACGAAATAATAAGAAAGAGATGATTTAA 296

DB 35 AACAGATCTCTGAAGAGCTAGACGAGTCTTACGAGCGCTTACGCGGAGACAGCGGG 94

QY 297 ACCAGAGAAACGCTTACAGCAGCTTCTCCAGAGAGCACTAATTAATAGTCAAGAAATGG 356

DB 95 CGCAGAACGCGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 154

QY 357 GAGATGAAAAATACAGATTTGTACACAAATGCTGGAATTTGGTGAATAATCGGCAAGAC 416

DB 155 GCGAGAGAGATCCAGATCTGTGAGCCAGATGTTGGAGTGTGTTGAGAACCGCAGCGGC 214

QY 417 AAATGGAGTTACACTCACAGTGTTCACAGATC----- 449

DB 215 AGTGGACACGACCGTGGAGCTGTTTCGAGGCGGCGCAGCAGGAGCTGGGCGACACAGTGGGCA 274

QY 450 -CTGCTGAAGTGAAGAGCGCTCAGATGAAGCAAGCAAGATGATTCAGCAACCAAGCAAGAA 508  
 Db 275 ACAGCGCAAGTGTGGCGGACAGCGCCCAATGCGCATGGGTAGCGAGTCTGTGACAGC 334  
 QY 509 TCTTCAAGAA--GACCCCGCAGCGGACGACAGTGAAGCCGTGATTTATGTACATGG 566  
 Db 335 CCAACAGCAAGCGCTCAGCGCGGACGCGCAACACAGCAAGACCGGTGAGACGCGTCCAGCA 394  
 QY 567 CAAATGGGATTGAAGACTGTGATGATCAGCCACCTAAAGAAAGAAATCCAGTCAGCAA 626  
 Db 395 ACCACGACAGCAGCGCGGCTCGGCGACACCCCAAGGAGAAAGGCGCAAGACCTCCA 454  
 QY 627 AGAAAAGAAAGCTCCAAAGCGGCAAGCAAGGAGGAGCTTCACCTGTTGAGTTTGCAA 686  
 Db 455 AGAAGAAGAGCGCTCCAAAGCGGCAAGCGGAGGAGCGGTCCCTCGGACCTCCCA 514  
 QY 687 TAGATCTTAATGAACCTACATAGTGTATGCAACCAAGTGTCTTATGGGAGATGATAG 746  
 Db 515 TCGACCCCAAGCAAGCGGAGTGTCTGTGCAACCAAGTGTCTTATGGGAGATGATCG 574  
 QY 747 GATGTGACATGAACAGTGTCCAAATGAATGGTTTTCATCTTTCATGTTTCACTTACCT 806  
 Db 575 GCTGCGACACAGCAGTGTCCCATCGAGTGTTCACCTCTCTGCTGGGCTCAATC 634  
 QY 807 ATAAACCAAGGGGAAATGTTATGTCACCAAGTGTCTTATGGGAGATGATAG 866  
 Db 635 ATAAACCAAGGGGAAATGTTATGTCACCAAGTGTCTTATGGGAGATGATAG 694  
 QY 867 ACAAAAGTACTGAAGAGCAAAAAGGATAG 897  
 Db 695 ACAAAAGCTGGAGAAATCCAAAAGAGAG 725

RESULT 14

AAV62285  
 ID AAV62285 standard; cDNA; 1902 BP.  
 AC XX AAV62285;  
 DT 18-JAN-1999 (first entry)  
 XX Partial INGI partial cDNA sequence.  
 DE INGI gene; p33ING1; human; apoptosis; cell death; breast cancer;  
 KW brain tumour; gene therapy; tumour suppressor; ss.  
 XX Homo sapiens.  
 XX Key Location/Qualifiers  
 FH CDS 109..741  
 FT /\*tag= a  
 XX WO9844102-A2.  
 XX 08-OCT-1998.  
 XX 26-MAR-1998; 98WO-CA00277.  
 XX 27-MAR-1997; 97US-0828158.  
 XX (UYTE-) UNIV TECHNOLOGIES INT INC.  
 PA Garkavtsev I, Helbing CC, Johnston RN, Rjabowol K;  
 PI WPI; 1998-542700/46.  
 XX P-PSDB; AAW79674.  
 DR Modulating eukaryotic apoptosis by increasing p33ING1 activity -  
 PT using p33ING1 derivatives, to induce apoptosis in cancer cells, and  
 PT in the investigation of apoptotic pathways  
 XX Example 2; Fig 2; 66pp; English.  
 PS  
 XX

CC This is the nucleotide sequence of a human INGI (inhibitor of  
 CC growth) partial cDNA clone that codes for a p33ING1 polypeptide  
 CC (see AAW79674), a novel inhibitor of cell growth and a candidate  
 CC tumour suppressor. INGI is a new gene that is expressed in normal  
 CC mammary epithelial cells, but which is expressed only at lower  
 CC levels in several cancerous mammary epithelial cell lines and is  
 CC not expressed in many primary brain tumours. To isolate INGI, a  
 CC subtractive hybridisation of breast cancer cell line cDNAs was  
 CC performed with cDNA from normal mammary epithelial cells, and  
 CC subcloned cDNAs were cloned into retrovirus vector pLNCX.  
 CC Following passage through a packaging line, normal mouse mammary  
 CC epithelial cells were infected, and infected cells were injected  
 CC into nude mice. Putative transforming fragments from tumours were  
 CC isolated by PCR (see AAV62290-91) and subcloned into LNCX. An INGI  
 CC fragment was obtained and used to screen normal human fibroblast  
 CC and HeLa cell cDNA libraries. 2 clones were sequenced to obtain  
 CC the partial INGI sequence. The complete cDNA sequence (see  
 CC AAV62292) was obtained by RACE. A claimed method to potentiate  
 CC apoptosis in a eukaryotic cell involves administering an active  
 CC p33ING1 peptide or an oligonucleotide encoding such as a peptide.  
 CC A claimed method for inhibiting apoptosis in a eukaryotic cell  
 CC involves administering an antisense oligonucleotide. Also claimed  
 CC are a method for determining the apoptotic characteristics of a  
 CC eukaryotic cell, an assay for determining the level of p33ING1  
 CC activity in a eukaryotic cell, and an isolated eukaryotic cell  
 CC substantially free of p33ING1 biological activity. The invention  
 CC discloses INGI derivatives or variants that may be used to induce  
 CC apoptosis in eukaryotic cancer cells.  
 XX  
 SQ Sequence 1902 BP; 574 A; 390 C; 462 G; 476 T; 0 other;  
 Query Match 17.0%; Score 183.8; DB 19; Length 1902;  
 Best Local Similarity 57.0%; Pred. No. 4.1e-32;  
 Matches 394; Conservative 0; Mismatches 267; Indels 30; Gaps 2;  
 QY 237 AAGAAACGTTAAGGAAATGTATGATGCTCTACGAAAAATATAGAAGAAGATGATTAA 296  
 Db 35 AACAGATCCTGAAGGAGCTAGACGAGTGTCTACGAGCGCTTCAGTCGCGACAGACGGG 94  
 QY 297 ACCAGAGAAAGCTCTACAGAGCTTCTCCAGAGCAGCTAATTAATAGTCAGAAATGG 356  
 Db 95 CGCAGAAAGCGCGGATGCTGCTGTCAGCGCGCGCTGATCCCGACGAGAGCTGG 154  
 QY 357 GAGATGAAAAATACAGATTTGTACACAAATGCTCGAATTTGGTGGAAATCGGGCAAGAC 416  
 Db 155 GCGACGAGAAGATCCAGATCGTCAGCCAGATGCTGAGCTGTGGAGAACCGCACCGGC 214  
 QY 417 AATGGAGTTACACTCACAGTGTTCCTCAAGATC----- 449  
 Db 215 AGGTGGACAGCCACGCTGGAGCTGTCGAGCGCGCAGCAGGAGCTGGCGCACACAGTGGCA 274  
 QY 450 -CTGCTGAAGTGAAGAGCGCTCAGATGAAGCAAGATGATTCAGCAACCAAGCAAGAA 508  
 Db 275 ACAGCGCAAGTGTGGCGGACAGCGCCCAATGCGCATGGGTAGCGAGTCTGTGACAGC 334  
 QY 509 TCTTCAAGAA--GACCCCGCAGCGGACGACAGTGAAGCCGTGATTTATGTACATGG 566  
 Db 335 CCAACAGCAAGCGCTCAGCGCGGACGCGCAACACAGCAAGACCGGTGAGACGCGTCCAGCA 394  
 QY 567 CAAATGGGATTGAAGACTGTGATGATCAGCCACCTAAAGAAAGAAATCCAGTCAGCAA 626  
 Db 395 ACCACGACAGCAGCGCGGCTCGGCGACACCCCAAGGAGAAAGGCGCAAGACCTCCA 454  
 QY 627 AGAAAAGAAAGCTCCAAAGCGGCAAGCAAGGAGGAGCTTCACCTGTTGAGTTTGCAA 686  
 Db 455 AGAAGAAGAGCGCTCCAAAGCGGCAAGCGGAGGAGCGGTCCCTCGGACCTCCCA 514  
 QY 687 TAGATCTTAATGAACCTACATAGTGTATGCAACCAAGTGTCTTATGGGAGATGATAG 746  
 Db 515 TCGACCCCAAGCAAGCGGAGTGTCTGTGCAACCAAGTGTCTTATGGGAGATGATCG 574  
 QY 747 GATGTGACATGAACAGTGTCCAAATGAATGGTTTTCATCTTTCATGTTTCACTTACCT 806



Db 575 GCTGGACACGACGAGTGTCCCATCGAGTGGTTCCTCCTGCTGGGTGGGTCAATC 634  
 Qy 807 ATAAACCAAGGGGAAATGGTATGTCGCAAGTGCAGGGGAGATAATGAGAAAACAATGG 866  
 Db 635 ATAAACCAAGGGGCAAGTGGTACTGTCCCAAGTGCCTGGGGGAGAACGAGACCATGG 694  
 Qy 867 ACAAAAGTACTGAAAGACAAACAAAAGGATAG 897  
 Db 695 ACAAGCCCTGGAGAAATCCAAAAGAGAG 725

RESULT 15  
 AAT69652  
 ID AAT69652 standard; cDNA; 2061 BP.  
 XX  
 AC AAT69652;  
 XX  
 DT 27-AUG-1997 (first entry)  
 XX  
 DE Tumour suppressor gene INGI full-length cDNA.  
 XX  
 KW Tumour suppressor gene; INGI; p33INGI; breast cancer; brain cancer;  
 KW diagnosis; gene therapy; ss.  
 XX  
 OS Homo sapiens.

XX Key Location/Qualifiers  
 FH CDS 16..900  
 FT /\*tag= a  
 FT primer\_bind complement (474..494)  
 FT /\*tag= b  
 FT /\*note= "direct primer 1"  
 FT complement (763..782)  
 FT /\*tag= c  
 FT primer\_bind complement (857..876)  
 FT /\*tag= d  
 FT /\*note= "reverse primer 3"  
 FT primer\_bind 890..900  
 FT /\*tag= e  
 FT /\*note= "reverse primer 4"

XX WO9721809-A1.  
 XX 19-JUN-1997.  
 XX 06-DEC-1996; 96WO-CA00819.  
 XX 15-NOV-1996; 96US-0751230.  
 XX 08-DEC-1995; 95US-0569721.  
 XX (UYTE-) UNIV TECHNOLOGIES INT INC.  
 XX Garkavstev I, Riabowol K;  
 XX WPI; 1997-332781/30.  
 XX P-PSDB; AAW18119.

Isolated tumour suppressor gene, INGI - useful to develop products for inhibiting or increasing cell proliferation, in particular for treatment or diagnosis of cancer

Claim 11; Fig 3; 63pp; English.  
 A full-length cDNA clone (AAT69652), designated INGI, codes for a novel 33 kDa tumour suppressor protein p33INGI (AAW19119), formerly p33igl, that is a potent inhibitor of cell growth. A partial clone (AAT69651) was isolated by subtractive hybridisation between normal mammary and transformed epithelial cDNAs, isolation of an antisense INGI cDNA insert that caused increased cell proliferation, and use of the insert to screen normal human fibroblast and HeLa cDNA libraries. The complete INGI sequence was then obt'd. by 5'RACE. INGI is localised to the 13q33-34 chromosome region, to which a

CC number of human cancers have been mapped. INGI nucleic acids can be used in the diagnosis of breast cancer; a decreased level of INGI mRNA indicates cancerous cells. They can also be used in gene therapy methods to block the proliferation of cancer cells.  
 XX  
 SQ Sequence 2061 BP; 602 A; 439 C; 515 G; 505 T; 0 other;  
 Query Match 17.0%; Score 183.8; DB 18; Length 2061;  
 Best Local Similarity 57.0%; Pred. No. 4.2e-32;  
 Matches 394; Conservative 0; Mismatches 267; Indels 30; Gaps 2;

Qy 237 AAGAAACGTTAAAGGAATTTGATGATCTACGAAAAATATAGAAAGATGATTTAA 296  
 Db 194 AACAGATCTGAAGGAGCTAGACGAGTGTCTAGGAGCGCTTACGTCGAGACAGCGGG 253  
 Qy 297 ACCAGAGAAACGCTCTACAGCAGCTTCTCCAGAGAGACACTAATTAATAGTCAAGATGG 356  
 Db 254 CGCAGAACGGGGGATGTGCACTGTGTGACGCGCGCTGATCCGAGCAGGAGCTGG 313  
 Qy 357 GAGATGAAAAATACAGATTGTTACACAAATGCTCGAATTTGGTGGAAATCGGGCAAGAC 416  
 Db 314 CGCAGGAGAAATCCAGATCGTGTGACCCAGATGTTGAGCTGTGTGAGAACCCGACGCG 373  
 Qy 417 AATGGAGTTACACTACAGTGTGTTCCAGATC----- 449  
 Db 374 AGGTGGACAGCCACGCTGGAGCTGTTTCGAGGCGCAGCAGGAGCTGGGCGACACAGTGGCA 433  
 Qy 450 -CTGCTGAAGTGAACGAGCGCTCAGATAAAGCAAGATGGATTCAGCCCAACAGCAAGAA 508  
 Db 434 ACAGCGCAAGTGTGGCGCGGACAGGCCCAATGGCGATGCGGTACGCGAGTGTGACAGC 493  
 Qy 509 TCTTCAAGAA--GACCCCGCAGCGACGAGCCAGTGAAGCGCGTATTATGTACATGG 566  
 Db 494 CCAACAGCAAGCGCTCACGGCGGAGCGCAACAGAGAACCGTGAGAAGCGGTCCAGCA 553  
 Qy 567 CAATGGGATTGAAGACTGTGATGATCAGCCACCTAAAGAAAGAAATCCAAAGTCAAGAA 626  
 Db 554 ACCAGCACCAGCAGCAGCGCGCTCGGGCACACACCCCAAGGAGAAAGGCCAAGACTCCA 613  
 Qy 627 AGAAAAAGAAAGCGTCCCAAGCGCAGGAGGAAAGGAAAGCTTACCTGTTGAGTTTGA 686  
 Db 614 AGAAGAAAGCGCTCCCAAGCGCAGGCGGAGGAGCGTCCCTGCCGACCTCCCA 673  
 Qy 687 TAGATCTTAATGAACCTACATACTACTTATGCAACCAAGTGTCTTATGGGAGATGATAG 746  
 Db 674 TCGACCCCAACGACCCACGTAATGATGTCGTCGCAACCCAGTCTCCTATGGGAGATGATCG 733  
 Qy 747 GATGTACAATGAACAGTGTCCAATGATGATGTTTCACTTTTCACTGTGTTTCACTTACT 806  
 Db 734 GCTGCGACACGACGAGTGCCTCATTCGAGTGGTTCCTTCTCTGCTGCGTGGGCTCAATC 793  
 Qy 807 ATAAACCAAGGGGAAATGGTATTCGCCAAAGTGCAGGGGAGATAATGAGAAAAACAATGG 866  
 Db 794 ATAAACCAAGGGCAAGTGGTACTCTCCCAAGTGCCTGGGGGAGAACGAGAACCATGG 853  
 Qy 867 ACAAAAGTACTGAAAGACAAACAAAAGGATAG 897  
 Db 854 ACAAGCCCTGGAGAAATCCAAAAGAGAG 884

Search completed: June 20, 2003, 01:27:32  
 Job time : 299 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2003, 01:19:25 ; Search time 2824 Seconds  
(without alignments)  
11129.968 Million cell updates/sec

Title: US-09-513-365a-2

Perfect score: 1080

Sequence: 1 ggcggccgcggcggtgcatg.....aaaaaaaaaaaaaaaaaaaa 1080

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl :

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.inu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg\_hum.\*
- 31: em.htg\_inv.\*
- 32: em.htg\_other.\*
- 33: em.htg\_mus.\*
- 34: em.htg\_pln.\*
- 35: em.htg\_rod.\*
- 36: em.htg\_man.\*
- 37: em.htg\_vrt.\*
- 38: em.sy.\*
- 39: em.htgo\_hum.\*
- 40: em.htgo\_mus.\*
- 41: em.htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1080	100.0	1080	6	AX211560	Sequence
2	1080	100.0	1080	9	AF053537	Homo sapi
3	1073	99.4	1141	9	BC030128	Homo sapi
4	1046	96.9	1078	9	AB012853	Homo sapi
5	839.8	77.8	843	9	HSJ06851	Homo sapi
6	817.2	75.7	974	9	HSJ06252	Homo sapi
7	817.2	75.7	189972	2	AC107214	Homo sapi
8	761.4	70.5	1001	10	AF078834	Mus muscu
9	686.8	63.6	168245	2	AC128348	Rattus no
c	574.6	53.2	42029	2	AC114762	Homo sapi
11	432.8	40.1	1094	5	AY014017	Xenopus l
12	273.6	25.3	188087	2	AC129658	Rattus no
13	256.8	23.8	840	10	AF149820	Mus muscu
14	253.6	23.5	1584	10	BC016573	Mus muscu
15	239.2	22.1	423	9	HSJ06251	Mus muscu
16	237.4	22.0	873	6	AR171884	Sequence
17	237.4	22.0	911	9	AF044076	Homo sapi
18	237.4	22.0	1533	9	AF149721	Homo sapi
19	237.4	22.0	2444	9	AB024401	Homo sapi
20	236	21.9	840	9	AF078835	Homo sapi
21	236	21.9	845	9	HSJ06251	Homo sapi
22	236	21.9	1533	6	AX367042	Sequence
23	236	21.9	2897	9	AF181850	Homo sapi
24	236	21.9	188087	2	AC129658	Rattus no
c	211.4	19.6	188087	2	AC129658	Rattus no
25	207	19.2	166219	2	AC120896	Mus muscu
c	199	18.4	235079	2	AC124475	Mus muscu
27	198.8	18.4	1959	10	AF177755	Mus muscu
28	197.8	18.3	1944	10	AF177756	Mus muscu
29	184.4	17.1	1980	9	AB024404	Homo sapi
30	183.8	17.0	1074	9	AF149722	Homo sapi
31	183.8	17.0	1143	6	AX367043	Sequence
32	183.8	17.0	1902	6	AR079046	Sequence
33	183.8	17.0	1902	6	AR087457	Sequence
34	183.8	17.0	1902	6	AR110646	Sequence
35	183.8	17.0	1902	6	AR154563	Sequence
36	183.8	17.0	2061	6	AR087464	Sequence
37	183.8	17.0	2061	6	AR110653	Sequence
38	183.8	17.0	2061	6	AR154570	Sequence
39	183.8	17.0	2061	6	AF001954	Homo sapi
40	183.8	17.0	2306	9	AB031269	Homo sapi
41	183.8	17.0	2891	9	AB024402	Homo sapi
42	183.8	17.0	1189	9	AF149723	Homo sapi
43	183.4	17.0	2096	9	HSP33ING2	Homo sapi
44	183.4	17.0	2552	9	HSJ06251	Homo sapi
45	183.4	17.0	2552	9	HSJ06251	Homo sapi

ALIGNMENTS

RESULT 1  
AX211560

LOCUS

DEFINITION Sequence 7 from Patent WO0159114.

ACCESSION AX211560

VERSION AX211560.1 GI:15523814

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS Harris,C.C. and Nagashima,M.

TITLE Tumour suppressor gene, p47ing3

JOURNAL Patent: WO 0159114-A 7 16-AUG-2001;

linear PAT 06-SEP-2001

THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)  
 Location/Qualifiers  
 1. .1080  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /note="p33ING2"  
 BASE COUNT 375 a 196 c 271 g 238 t  
 ORIGIN  
 Query Match 100.0%; Score 1080; DB:6; Length 1080;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-222;  
 Matches 1080; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCCGCGCGCGTGCATGTCGGCTGCTGATGCGGAGCGCGCGCGCGCGCGAT 60  
 Db 1 GCGGCCGCGCGCGTGCATGTCGGCTGCTGATGCGGAGCGCGCGCGCGCGAT 60  
 QY 61 CGGCAGGATGTAGGCGAGCAGCAGCAACTGTACTGTCGGCTGCGCTCTGACCGG 120  
 Db 61 CGGCAGGATGTAGGCGAGCAGCAGCAACTGTACTGTCGGCTGCGCTCTGACCGG 120  
 QY 121 GGAGCGAGCGCGCTGCTCACCTGCTACGTGCGAGGACTACCTTGAGTGGTGGAGTCGCT 180  
 Db 121 GGAGCGAGCGCGCTGCTCACCTGCTACGTGCGAGGACTACCTTGAGTGGTGGAGTCGCT 180  
 QY 181 GCGCCACGATGCGAGGAGCACTGCTGCTGCTGCGAGAGCTGGACAACAATATCAAGA 240  
 Db 181 GCGCCACGATGCGAGGAGCACTGCTGCTGCTGCGAGAGCTGGACAACAATATCAAGA 240  
 QY 241 AACGTTAAAGGAAATGTATGCTGCTACGAAATAATATAGAAAGAGATGATTTAAACCA 300  
 Db 241 AACGTTAAAGGAAATGTATGCTGCTACGAAATAATATAGAAAGAGATGATTTAAACCA 300  
 QY 301 GAAGAACTGCTACAGCAGCTTCTCCAGAGAGCACTAATTAATAGTCAAGAATGGGAGA 360  
 Db 301 GAAGAACTGCTACAGCAGCTTCTCCAGAGAGCACTAATTAATAGTCAAGAATGGGAGA 360  
 QY 361 TGAAGAAATACAGATGTTTACAAATGCTCGAATGTTGGTGGAAATCGGCAAGACAAAT 420  
 Db 361 TGAAGAAATACAGATGTTTACAAATGCTCGAATGTTGGTGGAAATCGGCAAGACAAAT 420  
 QY 421 GGAGTTACACTACAGTGTTCAGATCTGCTGAAAGTGAACGAGCTCAGATAAGC 480  
 Db 421 GGAGTTACACTACAGTGTTCAGATCTGCTGAAAGTGAACGAGCTCAGATAAGC 480  
 QY 481 AAAGATGGATTCAGCAACACCAAGAGATCTTCAAGAACACCCGAGCGAGCGACAG 540  
 Db 481 AAAGATGGATTCAGCAACACCAAGAGATCTTCAAGAACACCCGAGCGAGCGACAG 540  
 QY 541 TGAAGCGGTGATTTATGTCACATGCGCAATGGGATTTGAAGACTGTGTGATCAGCCACC 600  
 Db 541 TGAAGCGGTGATTTATGTCACATGCGCAATGGGATTTGAAGACTGTGTGATCAGCCACC 600  
 QY 601 TGAAGAAAGAAATCCAGTCAGCAAGAAAGAAAGAACGCTCCAGGCCAACGAGGAAG 660  
 Db 601 TGAAGAAAGAAATCCAGTCAGCAAGAAAGAAAGAACGCTCCAGGCCAACGAGGAAG 660  
 QY 661 GGAAGCTTCACCTGTTGAGTTGCAATGATGCTTAATGAACCTTACATGCTGTTATGCAA 720  
 Db 661 GGAAGCTTCACCTGTTGAGTTGCAATGATGCTTAATGAACCTTACATGCTGTTATGCAA 720  
 QY 721 CCAAGTGTCTTATGGGAGATGATAGGATGTGACAAATGAACAGTGTCCAAATGAATGGTT 780  
 Db 721 CCAAGTGTCTTATGGGAGATGATAGGATGTGACAAATGAACAGTGTCCAAATGAATGGTT 780  
 QY 781 TCACCTTTTCATGTTTTCACCTTACCTATTAACCAAGGGAATGATGTTGCCAAAGTG 840  
 Db 781 TCACCTTTTCATGTTTTCACCTTACCTATTAACCAAGGGAATGATGTTGCCAAAGTG 840  
 QY 841 CAGGGAGATATAGAAAAAATGAGCAAAAGTACTGAAAAAGACAAAAAGATAGAG 900  
 Db 841 CAGGGAGATATAGAAAAAATGAGCAAAAGTACTGAAAAAGACAAAAAGATAGAG 900

QY 901 ATCAGAGTAGTAAGGCCATCCACATTTAAAGGGTATTGTCCTTTTATATATTCGTT 960  
 Db 901 ATCAGAGTAGTAAGGCCATCCACATTTAAAGGGTATTGTCCTTTTATATATTCGTT 960  
 QY 961 TCGCTTTCAGAAATGTTTTAGGCTAAATGCATAGACTATGCAATATTTTATCATTA 1020  
 Db 961 TCGCTTTCAGAAATGTTTTAGGCTAAATGCATAGACTATGCAATATTTTATCATTA 1020  
 QY 1021 GTATTATGCTGTTATTTAAAGTGTGTTGCTTTGAAAAAAGAAAAAAGAAAAA 1080  
 Db 1021 GTATTATGCTGTTATTTAAAGTGTGTTGCTTTGAAAAAAGAAAAAAGAAAAA 1080

RESULT 2  
 AF053537 1080 bp mRNA linear PRI 15-AUG-2001  
 LOCUS Homo sapiens p33 (ING2) mRNA, complete cds.  
 DEFINITION AF053537  
 ACCESSION AF053537  
 VERSION AF053537.1 GI:9992837  
 KEYWORDS  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1080)  
 AUTHORS Nagashima, M., Shiseki, M., Miura, K., Hagiwara, K., Linke, S.P.,  
 Pedoux, R., Wang, X.W., Yokota, J., Riabowol, K. and Harris, C.C.  
 TITLE DNA damage-inducible gene p33ING2 negatively regulates cell  
 proliferation through acetylation of p53  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (17), 9671-9676 (2001)  
 MEDLINE 21396501  
 PUBMED 11481424  
 REFERENCE 2 (bases 1 to 1080)  
 AUTHORS Nagashima, M., Hagiwara, K., Minter, A.R. and Harris, C.C.  
 TITLE Direct Submission  
 JOURNAL Submitted (12-MAR-1998) Laboratory of Human Carcinogenesis,  
 National Cancer Institute, 37 Convent Drive Bldg. 37 Rm. 2C01,  
 Bethesda, MD 20892, USA

FEATURES  
 Location/Qualifiers  
 1. .1080  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="4"  
 /map="4q35"  
 1. .1080  
 /gene="ING2"  
 68..910  
 /gene="ING2"  
 /codon\_start=1  
 /product="p33"  
 /protein\_id="AAG11395.1"  
 /db\_xref="GI:9992837"  
 /translation="MLGQQQOLYSALLTGERSLLTCYVDYLECVELPHDMQR  
 NYSVLRELNKKYQETLKEIDVYKKEKEDDLNKKRLQQLQRLINQELGDEKIQ  
 IYVLMLELVARAROMELHSQCQDPAESERASDKAMDSSQPERSSRRPRORTSES  
 RLCHMANGIEDCDQPPKPKSKAKKSKAKQERASPVFEAIDNPTCYCLN  
 QVSYGEMIGCDNEQCPLEWFHFSCVSLYKPKKWKYCPKRGDNKMDKTEKTKD  
 RSR"  
 106  
 variation  
 /gene="ING2"  
 /note="polymorphism"  
 /replaces="c"  
 375 a 196 c 271 g 238 t  
 BASE COUNT 375 a 196 c 271 g 238 t  
 ORIGIN  
 Query Match 100.0%; Score 1080; DB:9; Length 1080;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-222;  
 Matches 1080; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCCGCGCGCGTGCATGTCGGCTGCTGATGCGGAGCGCGCGCGCGCGAT 60  
 Db 1 GCGGCCGCGCGCGTGCATGTCGGCTGCTGATGCGGAGCGCGCGCGCGCGAT 60  
 QY 61 CGGCAGGATGTAGGCGAGCAGCAGCAACTGTACTGTCGGCTGCGCTCTGACCGG 120  
 Db 61 CGGCAGGATGTAGGCGAGCAGCAGCAACTGTACTGTCGGCTGCGCTCTGACCGG 120  
 QY 121 GGAGCGAGCGCGCTGCTCACCTGCTACGTGCGAGGACTACCTTGAGTGGTGGAGTCGCT 180  
 Db 121 GGAGCGAGCGCGCTGCTCACCTGCTACGTGCGAGGACTACCTTGAGTGGTGGAGTCGCT 180  
 QY 181 GCGCCACGATGCGAGGAGCACTGCTGCTGCTGCGAGAGCTGGACAACAATATCAAGA 240  
 Db 181 GCGCCACGATGCGAGGAGCACTGCTGCTGCTGCGAGAGCTGGACAACAATATCAAGA 240  
 QY 241 AACGTTAAAGGAAATGTATGCTGCTACGAAATAATATAGAAAGAGATGATTTAAACCA 300  
 Db 241 AACGTTAAAGGAAATGTATGCTGCTACGAAATAATATAGAAAGAGATGATTTAAACCA 300  
 QY 301 GAAGAACTGCTACAGCAGCTTCTCCAGAGAGCACTAATTAATAGTCAAGAATGGGAGA 360  
 Db 301 GAAGAACTGCTACAGCAGCTTCTCCAGAGAGCACTAATTAATAGTCAAGAATGGGAGA 360  
 QY 361 TGAAGAAATACAGATGTTTACAAATGCTCGAATGTTGGTGGAAATCGGCAAGACAAAT 420  
 Db 361 TGAAGAAATACAGATGTTTACAAATGCTCGAATGTTGGTGGAAATCGGCAAGACAAAT 420  
 QY 421 GGAGTTACACTACAGTGTTCAGATCTGCTGAAAGTGAACGAGCTCAGATAAGC 480  
 Db 421 GGAGTTACACTACAGTGTTCAGATCTGCTGAAAGTGAACGAGCTCAGATAAGC 480  
 QY 481 AAAGATGGATTCAGCAACACCAAGAGATCTTCAAGAACACCCGAGCGAGCGACAG 540  
 Db 481 AAAGATGGATTCAGCAACACCAAGAGATCTTCAAGAACACCCGAGCGAGCGACAG 540  
 QY 541 TGAAGCGGTGATTTATGTCACATGCGCAATGGGATTTGAAGACTGTGTGATCAGCCACC 600  
 Db 541 TGAAGCGGTGATTTATGTCACATGCGCAATGGGATTTGAAGACTGTGTGATCAGCCACC 600  
 QY 601 TGAAGAAAGAAATCCAGTCAGCAAGAAAGAAAGAACGCTCCAGGCCAACGAGGAAG 660  
 Db 601 TGAAGAAAGAAATCCAGTCAGCAAGAAAGAAAGAACGCTCCAGGCCAACGAGGAAG 660  
 QY 661 GGAAGCTTCACCTGTTGAGTTGCAATGATGCTTAATGAACCTTACATGCTGTTATGCAA 720  
 Db 661 GGAAGCTTCACCTGTTGAGTTGCAATGATGCTTAATGAACCTTACATGCTGTTATGCAA 720  
 QY 721 CCAAGTGTCTTATGGGAGATGATAGGATGTGACAAATGAACAGTGTCCAAATGAATGGTT 780  
 Db 721 CCAAGTGTCTTATGGGAGATGATAGGATGTGACAAATGAACAGTGTCCAAATGAATGGTT 780  
 QY 781 TCACCTTTTCATGTTTTCACCTTACCTATTAACCAAGGGAATGATGTTGCCAAAGTG 840  
 Db 781 TCACCTTTTCATGTTTTCACCTTACCTATTAACCAAGGGAATGATGTTGCCAAAGTG 840  
 QY 841 CAGGGAGATATAGAAAAAATGAGCAAAAGTACTGAAAAAGACAAAAAGATAGAG 900  
 Db 841 CAGGGAGATATAGAAAAAATGAGCAAAAGTACTGAAAAAGACAAAAAGATAGAG 900

[illegible]

ACCESSION	VERSION	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	REMARK	COMMENT
MGC:10524	IMAGE:3941655	mrna	complete cds.						
BC030128	1	GI:20987212							
MGC									
Homo sapiens									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
1 (bases 1 to 1141)									
Direct Submission									
Submitted (07-MAY-2002)									
National Institutes of Health, Mammalian									
Gene Collection (MGC), Cancer Genomics Office, National Cancer									
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,									
USA									
NIH-MGC Project URL: http://mgc.nci.nih.gov									
Contact: MGC help desk									
Email: cgabs-r@mail.nih.gov									
Tissue Procurement: DCTD/DTP									
cDNA Library Preparation: Rubin Laboratory									
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)									
DNA Sequencing by: National Institutes of Health Intramural									
Sequencing Center (NISC),									
Gaithersburg, Maryland;									
Web site: http://www.nisc.nih.gov/									
Contact: nisc.mgc@hri.nih.gov									
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,									
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,									
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,									
Hansen,N., Ho,S.-L., Karlins,E., Laric,P., Legaspi,R., Maduro,Q.L.,									
Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J.,									
Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W., Tsurgeon,C.,									
Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A.,									
Zhang,L.-H. and Green,E.D.									
Clone distribution: MGC clone distribution information can be found									
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov									
Series: IRAL Plate: 15 Row: m Column: 11									
This clone was selected for full length sequencing because it									
passed the following selection criteria: matched mRNA gi: 4504694.									
Location/Qualifiers									
1. 1141									
/organism="Homo sapiens"									
/db_xref="LocusID:3622"									
/db_xref="taxon:9606"									
/clone="MGC:10524 IMAGE:3941655"									
/tissue_type="Lung, small cell carcinoma"									
/clone_lib="NIH_MGC_7"									
/lab_host="DH10B-R"									
/note="Vector: pOTB7"									
136. 978									
/codon_start=1									
/product="inhibitor of growth family, member 1-like"									
/protein_id="AA30128.1"									
/db_xref="GI:20987213"									
/translation="MLGQQQQQLLYSAAALLTGRSRLTYVDYLYECVSLPHDMQR									
IVSVLRELNLQKETLEIKEDVYKYKEDDLNQKKRLQLQRLALINSQELGDKTQ									
IVTOMELVENAROMELSHQCFQPAESERASDQKAKMDSQPERSSRRRPRRTQS									
ROLCHMANGIECDQDPPEKESKSAKKRKAQERASPVFEALDPNPTCYCLCN									
QYSYGMICGDNEQCPIEWFHFSVSLYTPKQWYCPKRGDNEKTMDSKSTKTKND									
RRSR"									
BASE COUNT	373 a	221 c	302 g	245 t					
ORIGIN									
Query Match	99.4%;	Score 1073;	DB 9;	Length 1141;					
Best Local Similarity	100.0%;	Pred. No. 1e-220;							
Matches 1073;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;	
QY	1	CGCGCGCGCGCGTGCATGTGCGGCTGCTGGATCGGAGCGCGCGCGCGCGGAT	60						
Db	69	CGCGCGCGCGCGTGCATGTGCGGCTGCTGGATCGGAGCGCGCGCGCGCGGAT							

```
Db 129 CGCAGAGTGTAGGGCAGCAGCAGCAACTGTACTGCTGCGCTCTCTGACCG 188
QY 121 GGAGCGAGCGGCTGCTCACCCTGCTACGTGCAGGACTACCTTGAGTGGTGGAGTCGCT 180
Db 189 GGAGCGAGCGGCTGCTCACCCTGCTACGTGCAGGACTACCTTGAGTGGTGGAGTCGCT 248
QY 181 GCCCAGCAGATCAGAGGAACGTGCTGCTGCGAGAGCTGGCAACAATAATATCAAGA 240
Db 249 GCCCAGCAGATCAGAGGAACGTGCTGCTGCGAGAGCTGGCAACAATAATATCAAGA 308
QY 241 AAGCTTAAAGGAATGTATGATGCTACGAAATAATAAGAAAGAGATGATTTAAACCA 300
Db 309 AAGCTTAAAGGAATGTATGATGCTACGAAATAATAAGAAAGAGATGATTTAAACCA 368
QY 301 GAAGAAACCTTACAGCAGCTTCTCCAGAGAGCACTAATAATAGTCAAGAATGGGAGA 360
Db 369 GAAGAAACCTTACAGCAGCTTCTCCAGAGAGCACTAATAATAGTCAAGAATGGGAGA 428
QY 361 TGAATAAATACAGATTGTTACAAATGCTCGAATTTGGTGGAAATCGGGCAAGACAAAT 420
Db 429 TGAATAAATACAGATTGTTACAAATGCTCGAATTTGGTGGAAATCGGGCAAGACAAAT 488
QY 421 GGATTTACTACAGTGTGTTTCAAGATCCTGCTGAAAGTGAACGAGCCTCAGATAAAGC 480
Db 489 GGATTTACTACAGTGTGTTTCAAGATCCTGCTGAAAGTGAACGAGCCTCAGATAAAGC 548
QY 481 AAGATGGGATTCAGCAGCAACCAAGAGATCTTCAAGAACACCCCGCAGCGGACCAAG 540
Db 549 AAGATGGGATTCAGCAGCAACCAAGAGATCTTCAAGAACACCCCGCAGCGGACCAAG 608
QY 541 TGAAGCCGTGATTTATGTACATGCAATGCAATGGGATTTGAAGCTGTGATGATCAGCACCC 600
Db 609 TGAAGCCGTGATTTATGTACATGCAATGGGATTTGAAGCTGTGATGATCAGCACCC 668
QY 601 TAAGAAAGAAATCCAGTTCAGCAAGAAAGAAACGCTCAAGGCCAAGCAGGAAG 660
Db 669 TAAGAAAGAAATCCAGTTCAGCAAGAAAGAAACGCTCAAGGCCAAGCAGGAAG 728
QY 661 GGAAGCTTCACCTGTTGAGTTTGCATAGATCTTAATGAACCTACATACCTGCTTATGCA 720
Db 729 GGAAGCTTCACCTGTTGAGTTTGCATAGATCTTAATGAACCTACATACCTGCTTATGCA 788
QY 721 CCAAGTGTCTTATGGGAGATGATAGGATGTGACATGAACAGTGTCCCAATTTGAATGGTT 780
Db 789 CCAAGTGTCTTATGGGAGATGATAGGATGTGACATGAACAGTGTCCCAATTTGAATGGTT 848
QY 781 TCACCTTTTCATGTTTACCTTACCTTAAACCAAGGGGAATGGTATTTGCCCAAGTG 840
Db 849 TCACCTTTTCATGTTTACCTTACCTTAAACCAAGGGGAATGGTATTTGCCCAAGTG 908
QY 841 CAGGGAGAGATAATGAGAAACCAATGACAAAGTACTGAAAGACAAAAAGGATAGAAG 900
Db 909 CAGGGAGAGATAATGAGAAACCAATGACAAAGTACTGAAAGACAAAAAGGATAGAAG 968
QY 901 ATCGAGGTAGTAAAGCCATCCACATTTTAAAGGGTATTTGCTTTTATATAATTCGTT 960
Db 969 ATCGAGGTAGTAAAGCCATCCACATTTTAAAGGGTATTTGCTTTTATATAATTCGTT 1028
QY 961 TCGTTTTCAGAAATGTTTATAGGTAAATGCATGAAGCTATGCAATTAATTTTAAATCATTA 1020
Db 1029 TCGTTTTCAGAAATGTTTATAGGTAAATGCATGAAGCTATGCAATTAATTTTAAATCATTA 1088
QY 1021 GTATTAAATGCTTATTAAGTGTGTTGACTTTGAAATTTGAAATTTGAAATTTGAAATTT 1073
Db 1089 GTATTAAATGCTTATTAAGTGTGTTGACTTTGAAATTTGAAATTTGAAATTTGAAATTT 1141
```

## RESULT 4

```
AB012853
LOCUS      AB012853
DEFINITION Homo sapiens INGIL mRNA for INGILp, complete cds.
ACCESSION AB012853
          1078 bp      mRNA      linear      PRI 18-MAR-1999
```

```
AB012853.1 GI:4115554
INGILp.
Homo sapiens fetal tissue_lib:fetal-brain cDNA to mRNA.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
Shimada,Y., Saito,A., Suzuki,M., Takahashi,E. and Horie,M.
Cloning of a novel gene (INGIL) homologous to ING1, a candidate
tumor suppressor
Cytogenet. Cell Genet. 83 (3-4), 232-235 (1998)
99172097
2 (bases 1 to 1078)
Shimada,Y., Saito,A. and Horie,M.
Direct Submission
Submitted (07-APR-1998) to F. Cantu
Otsuka GEN Research Institute; Kagasuno, Kawauchi-cho,
Tokushima, Tokushima 771-0192, Japan
(E-mail:shim@otsuka.genome.ad.jp, Tel:81-886-65-2888(ex.2476),
Fax:81-886-37-1035)
Sequence updated (17-Apr-1998).
Location/Qualifiers
1..1078
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/map="4q35.1"
/dev_stage="fetal"
/tissue_lib="fetal-brain"
1..1078
/gene="INGIL"
92..934
/gene="INGIL"
/codon_start=1
/product="INGILp"
/protein_id="BAA36419.1"
/db_xref="GI:411555"
/translation="MLGQOQOOLYSAAALLTGERGRLLTCYVQDYLECVELPHDMOR
NYSVLRLDNKYETLKEIDDDVYKYEKKEDDLNOKRLOQLORALINSQELGDKIO
IYTOLELVENARQMEHLSCQFQDPAESERASDKAKMDSSQPSRRSRPRQRISSE
RDLCHMANGIEDCDQPPKRAKSAKAKRAKQREASPEFAIDNPEPTICLN
QVSYGMGICDNEQCIPIEMFHFCSVLTYPKPKWYCPKCRGDNKMDKTEKTKKD
RRSR"
BASE COUNT 356 a 206 c 280 g 236 t
ORIGIN
Query Match 96.9%; Score 1046; DB 9; Length 1078;
Best Local Similarity 99.5%; Pred. No. 6.5e-215;
Matches 1049; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 GCGGCGCGCGCGGTCATGTGCGGTGCTGATGCGGAGCGGCGGCGGCGGCGGAT 60
Db 25 GCGGCGCGCGCGGTCATGTGCGGTGCTGATGCGGAGCGGCGGCGGCGGCGGAT 84
QY 61 CGGCAGGATGTTAGGCGAGCAGCAGCAACTGCTGCTGCGGTGCGCTCCTGACCGG 120
Db 85 CGGCAGGATGTTAGGCGAGCAGCAGCAACTGCTGCTGCGGCGGCTCCTGACCGG 144
QY 121 GGAGCGGAGCGGCTGCTACCTGCTGCTGAGGAGTACCTGAGTGGTGGAGTCGCT 180
Db 145 GGAGCGGAGCGGCTGCTACCTGCTGCTGAGGAGTACCTGAGTGGTGGAGTCGCT 204
QY 181 GCCCAGCAGATCAGAGGAACGTGCTGCTGCGAGAGCTGGCAACAATAATATCAAGA 240
Db 205 GCCCAGCAGATCAGAGGAACGTGCTGCTGCGAGAGCTGGCAACAATAATATCAAGA 264
QY 241 AAGCTTAAAGGAATGTATGATGCTACGAAATAATAAGAAAGAGATGATTTAAACCA 300
Db 265 AAGCTTAAAGGAATGTATGATGCTACGAAATAATAAGAAAGAGATGATTTAAACCA 324
QY 301 GAAGAAACCTTACAGCAGCTTCTCCAGAGAGCACTAATAATAGTCAAGAATGGGAGA 360
Db 325 GAAGAAACCTTACAGCAGCTTCTCCAGAGAGCACTAATAATAGTCAAGAATGGGAGA 384
```



Db	721	TCATGTGTTTCACTTACCTATAACCAAGGGGAAATGGTATTGCCCAAGTGCAGGGA	780
Qy	848	GATAANTGAGAAAAACAATGGACAAAAGPACTGAAAAGACAAAAAAGGATAGAAGATCGAGG	907
Db	781	GATAATGAGAAAAACAATGGACAAAAGTACTGAAAAGACAAAAAAGGATAGAAGATCGAGG	840
Qy	908	TAG 910	
Db	841	TAG 843	
RESULT 6			
HSING2S2			
LOCUS			
DEFINITION	HSING2S2	974 bp DNA linear	PRI 08-SEP-2000
ACCESSION	Homo sapiens p33 (ING2) gene, exon 2, complete cds.		
VERSION	AF062748		
KEYWORDS	AF062748.1 GI:9992841		
SEGMENT	2 of 2		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 974)		
JOURNAL	Nagashima, M., Hagiwara, K., Minter, A.R. and Harris, C.C.		
FEATURES	Submitted (01-MAY-1998) Laboratory of Human Carcinogenesis,		
source	National Cancer Institute, 37 Convent Dr. Bldg. 37, Rm. 2C26,		
gene	Bethesda, MD 20892, USA		
mRNA	Location/Qualifiers		
CDS	1..974		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	order(AF062747.1:1..423,1..938)		
	/gene="ING2"		
	join(AF062747.1:1..239,124..938)		
	/gene="ING2"		
	/product="p33 ING2"		
	join(AF062747.1:68..239,124..794)		
	/gene="ING2"		
	/codon_start=1		
	/product="p33"		
	/protein_id="AAG11396.1"		
	/db_xref="GI:9992842"		
	/translation="MLGQQOOLYSSAALLTGPERSLLTCYVDYLECVESLPHDMOR		
	NVSVRELNDKNYQETLKEIDDDYKYKKEDDLNKKRLQQLLQRLINSELGDEKIQ		
	IVTQMLELVENRARQELHSHSQCFDPAESERASDKAMDSQPERSSRRPRORTSES		
	RDICHWANGIEDDDGPPPEKKSKAKKKRKAQERASPVFEAIDNPETTCICLN		
	QVSGYMGICGDNQCPIEFHFSCVSLTYKPKGKWKYCPCRGDNEKTMDSKTEKTKK		
	RRSR"		
	<1..123		
	/gene="ING2"		
	/number=1		
	124..938		
	/gene="ING2"		
	/number=2		
	918..923		
	/gene="ING2"		
	/evidence=experimental		
BASE COUNT	348 a 156 c 205 g 265 t		
ORIGIN			
Query Match	75.7%;	Score 817.2;	DB 9; Length 974;
Best Local Similarity	99.68;	Pred. No. 1.2e-165;	
Matches 819; Conservative	0; Mismatches 3; Indels 0; Gaps 0;		
Qy	233	TATCAAGAAACGGTTAAAGGAAATGTGATGTCTACGAAAAATATAAGAAAGATGAT	292
Db	117	TTTTTAGAACGGTTAAAGGAATGTGATGTCTACGAAAAATATAGAAAGATGAT	176
Qy	293	TTAAACCAGAAAGAACGTCTACAGACGCTTCTCCAGAGAGCACATAATTATAGTCAAGAA	352



REFERENCE  
AUTHORS  
TITLE  
JOURNAL

MO 63108, USA  
3 (bases 1 to 189972)  
Waterston, R. H.  
Direct Submission  
Submitted (10-APR-2002) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
On Apr 10, 2002 this sequence version replaced gi:18677601.

COMMENT

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu/gsc/index.shtml  
Contact: submissions@wustl.wustl.edu  
----- Project Information -----  
Center project name: H\_NH0367N14  
----- Summary Statistics -----  
Sequencing vector: M13; 0%  
Sequencing vector: plasmid; 100%  
Chemistry: Dye-terminator ET; 0% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 189941 bases at least Q40  
Consensus quality: 190973 bases at least Q30  
Consensus quality: 191699 bases at least Q20  
Insert size: 204000; agarose-fp  
Insert size: 195974; sum-of-contigs  
Quality coverage: 11.39 in Q20 bases; agarose-fp  
Quality coverage: 10.96 in Q20 bases; sum-of-contigs  
----- NOTE: This is a 'working draft' sequence. It currently  
\* consists of 8 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 1253: contig of 1252 bp in length  
\* 1253: gap of unknown length  
\* 1353: contig of 1847 bp in length  
\* 3193: gap of unknown length  
\* 3299: gap of unknown length  
\* 3300: contig of 1121 bp in length  
\* 4421: gap of unknown length  
\* 4521: contig of 1138 bp in length  
\* 5658: gap of unknown length  
\* 5758: gap of unknown length  
\* 5759: contig of 1108 bp in length  
\* 6866: gap of unknown length  
\* 6967: contig of 30151 bp in length  
\* 37117: gap of unknown length  
\* 37218: contig of 64185 bp in length  
\* 101403: gap of unknown length  
\* 101503: contig of 88470 bp in length.  
Location/Qualifiers  
1. 189972  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="4"  
/clone="RP11-367N14"  
1. 1252  
/note="assembly\_name:Contig21"  
1353. 3199  
/note="assembly\_name:Contig29"  
3300. 4420  
/note="assembly\_name:Contig44"  
4521. 5658  
/note="assembly\_name:Contig56"  
5759. 6866  
/note="assembly\_name:Contig58"  
6967. 37117  
/note="assembly\_name:Contig65"  
37218. 101402

/note="assembly\_name:Contig66"  
clone\_end:SP6  
vector\_side:left"  
101503. 189972  
/note="assembly\_name:Contig67"  
clone\_end:T7  
vector\_side:left"  
BASE COUNT 52803 a 40308 c 41211 g 54950 t 700 others  
ORIGIN

Query Match 75.7%; Score 817.2; DB 2; Length 189972;  
Best Local Similarity 99.6%; Pred. No. 3.8e-165;  
Matches 819; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 233 TATCAAGAAACGTTAAAGGAAATTTGATGATGCTTACGAAATAATAAGAAAGAGATGAT 292  
Db 146705 TTTTGTAGAACGTTAAAGGAAATTTGATGATGCTTACGAAATAATAAGAAAGAGATGAT 146764  
QY 293 TTAACCCAGAGAAACGCTTACAGCAGCTTTCTCCAGAGAGACACTTAATTAATAGTCAAGAA 352  
Db 146765 TTAACCCAGAGAAACGCTTACAGCAGCTTTCTCCAGAGAGACACTTAATTAATAGTCAAGAA 146824  
QY 353 TTGGGAGATCAAAAATACAGATTTTACACAATGCTCGAATTTGGTGGAAATCGGGCA 412  
Db 146825 TTGGGAGATCAAAAATACAGATTTTACACAATGCTCGAATTTGGTGGAAATCGGGCA 146884  
QY 413 AGACAAATGGAGTTTACACTCACAGTGTTCCTCAAGATCTCTGTAAGTGAACGAGCCTCA 472  
Db 146885 AGACAAATGGAGTTTACACTCACAGTGTTCCTCAAGATCTCTGTAAGTGAACGAGCCTCA 146944  
QY 473 GATAAAGCAAGATGGATTCCAGCCAAACCAAGAAAGATCTTCAAGAGAACCCCGCAGCAG 532  
Db 146945 GATAAAGCAAGATGGATTCCAGCCAAACCAAGAAAGATCTTCAAGAGAACCCCGCAGCAG 147004  
QY 533 CGGACCAAGTGAAGCCCGTATTATGTCACATGCCAATGGGATGGAAGCTGTGATGAT 592  
Db 147005 CGGACCAAGTGAAGCCCGTATTATGTCACATGCCAATGGGATGGAAGCTGTGATGAT 147064  
QY 593 CAGCACCTAAAGAAAGAAATCCAAAGTCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 652  
Db 147065 CAGCACCTAAAGAAAGAAATCCAAAGTCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 147124  
QY 653 CAGGAAAGGAGAGCTTCCCTGTTGAGTTTGCAATAGATCCTAATGAACCTACATATGTC 712  
Db 147125 CAGGAAAGGAGAGCTTCCCTGTTGAGTTTGCAATAGATCCTAATGAACCTACATATGTC 147184  
QY 713 TTATGCAACCAAGTGTCTTATGGGAGATGATGATGATGATGATGATGATGATGATGATGAT 772  
Db 147185 TTATGCAACCAAGTGTCTTATGGGAGATGATGATGATGATGATGATGATGATGATGATGAT 147244  
QY 773 GAATGGTTTCACTTTTTCATGTTTCACTTACCTTAAACCAAGGAGAAATGATGATGATGATGAT 832  
Db 147245 GAATGGTTTCACTTTTTCATGTTTCACTTACCTTAAACCAAGGAGAAATGATGATGATGAT 147304  
QY 833 CCAAAGTGCAGGGGAGAGATAATGAGAAACCAATGACAAAAGTACTGAAAGACAAAAAAG 892  
Db 147305 CCAAAGTGCAGGGGAGAGATAATGAGAAACCAATGACAAAAGTACTGAAAGACAAAAAAG 147364  
QY 893 GATAGAAGATCGAGGTAGTAAAGGCCCATCCACATTTTAAAGGGTTATTGTCCTTTTATAT 952  
Db 147365 GATAGAAGATCGAGGTAGTAAAGGCCCATCCACATTTTAAAGGGTTATTGTCCTTTTATAT 147424  
QY 953 AATTCGTTTGTCTTTCAGAAAATGTTTATAGGTAATGCATAAGACTATGCAATAATTTT 1012  
Db 147425 AATTCGTTTGTCTTTCAGAAAATGTTTATAGGTAATGCATAAGACTATGCAATAATTTT 147484  
QY 1013 AATCATTAGTATTAAATGGTGTATTAAAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1054  
Db 147485 AATCATTAGTATTAAATGGTGTATTAAAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 147526

RESULT 8  
AF078834

[illegible]

Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
 Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,  
 Miner,G., Milner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
 Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogun,M., Okwuonu,G.,  
 Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
 Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G.,  
 Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I.,  
 Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,  
 Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
 Tansey,N., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
 Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,  
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
 Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,  
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
 Weinstock,G., and Gibbs,R.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 168245)  
 Worley,K.C.  
 Direct Submission  
 Submitted (19-JUL-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
 ----- Project Information  
 Center project name: GZCX  
 Center clone name: CH230-129P3  
 ----- Summary Statistics  
 Sequencing vector: Plasmid;  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 115858 bases at least Q40  
 Consensus quality: 122576 bases at least Q30  
 Consensus quality: 127679 bases at least Q20  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 50 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1 1084: contig of 1084 bp in length  
 \* 1085 1184: gap of unknown length  
 \* 1185 2200: contig of 1016 bp in length  
 \* 2201 2300: gap of unknown length  
 \* 2301 3302: contig of 1002 bp in length  
 \* 3303 3402: gap of unknown length  
 \* 3403 4508: contig of 1106 bp in length  
 \* 4509 4608: gap of unknown length  
 \* 4609 5989: contig of 1381 bp in length  
 \* 5990 6089: gap of unknown length  
 \* 6090 7786: contig of 1696 bp in length  
 \* 7786 9044: gap of unknown length  
 \* 9045 9145: contig of 1159 bp in length  
 \* 9145 10325: contig of 1181 bp in length  
 \* 10326 11588: contig of 1163 bp in length  
 \* 11589 11688: gap of unknown length  
 \* 11689 13187: contig of 1499 bp in length  
 \* 13188 13287: gap of unknown length  
 \* 13288 14806: contig of 1519 bp in length  
 \* 14807 14906: gap of unknown length  
 \* 15982: contig of 1076 bp in length  
 \* 15983 16082: gap of unknown length  
 \* 16083 18040: contig of 1958 bp in length  
 \* 18041 18140: gap of unknown length  
 \* 18141 19750: contig of 1610 bp in length  
 \* 19751 21035: contig of 1185 bp in length  
 \* 21036 21135: gap of unknown length  
 \* 21136 22589: contig of 1454 bp in length  
 \* 22590 22689: gap of unknown length  
 \* 22690 23848: contig of 1159 bp in length  
 \* 23849 25639: contig of 1691 bp in length  
 \* 25640 25739: gap of unknown length  
 \* 25740 28181: contig of 2442 bp in length  
 \* 28182 28281: gap of unknown length  
 \* 28282 29875: contig of 1594 bp in length  
 \* 29876 32032: contig of 2057 bp in length  
 \* 32033 32132: gap of unknown length  
 \* 32133 33817: contig of 1585 bp in length  
 \* 33818 35711: contig of 1894 bp in length  
 \* 35712 35811: gap of unknown length  
 \* 35812 37778: contig of 1967 bp in length  
 \* 37779 37878: gap of unknown length  
 \* 37879 40213: contig of 2335 bp in length  
 \* 40214 40313: gap of unknown length  
 \* 40314 42827: contig of 2514 bp in length  
 \* 42828 42927: gap of unknown length  
 \* 42928 45270: contig of 2343 bp in length  
 \* 45271 45370: gap of unknown length  
 \* 45371 48107: contig of 2737 bp in length  
 \* 48108 48207: gap of unknown length  
 \* 48209 51184: contig of 2977 bp in length  
 \* 51185 51284: gap of unknown length  
 \* 51285 53393: contig of 2109 bp in length  
 \* 53394 53493: gap of unknown length  
 \* 53494 56791: contig of 3298 bp in length  
 \* 56792 56891: gap of unknown length  
 \* 56892 62150: contig of 5259 bp in length  
 \* 62151 62250: gap of unknown length  
 \* 62251 65330: contig of 3080 bp in length  
 \* 65331 65430: gap of unknown length  
 \* 65431 69988: contig of 4468 bp in length  
 \* 69989 74110: contig of 4112 bp in length  
 \* 74111 74210: gap of unknown length  
 \* 74211 76637: contig of 2427 bp in length  
 \* 76638 82625: contig of 5888 bp in length  
 \* 82626 82725: gap of unknown length  
 \* 82726 85396: contig of 2671 bp in length  
 \* 85397 85496: gap of unknown length  
 \* 85497 89571: contig of 4075 bp in length  
 \* 89572 89671: gap of unknown length  
 \* 89672 94810: contig of 5139 bp in length  
 \* 94811 94910: gap of unknown length  
 \* 94911 98354: contig of 3444 bp in length  
 \* 98355 98454: gap of unknown length  
 \* 98455 103316: contig of 4862 bp in length  
 \* 103317 103416: gap of unknown length  
 \* 103417 108022: contig of 4606 bp in length  
 \* 108023 108122: gap of unknown length  
 \* 108123 114307: contig of 6185 bp in length  
 \* 114308 114407: gap of unknown length  
 \* 114408 121007: contig of 6600 bp in length  
 \* 121009 121107: gap of unknown length  
 \* 121108 127326: contig of 6219 bp in length  
 \* 127327 127426: gap of unknown length  
 \* 127427 136654: contig of 9228 bp in length  
 \* 136655 136754: gap of unknown length  
 \* 136755 148193: contig of 11439 bp in length

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

```
* 148194 148293: gap of unknown length
* 148294 157304: contig of 9011 bp in length
* 157305 157404: gap of unknown length
* 157405 168245: contig of 10841 bp in length.
FEATURES
    source      Location/Qualifiers
                1..168245
                /organism="Rattus norvegicus"
                /db_xref="taxon:10116"
                /clone="CH230-129p3"
BASE COUNT  43979 a 35809 c 36510 g 44464 t 7483 others
ORIGIN
Query Match      63.6%; Score 686.8; DB 2; Length 168245;
Best Local Similarity 90.5%; Pred. No. 4.4e-137;
Matches 744; Conservative 0; Mismatches 77; Indels 1; Gaps 1;
QY 233 TATCAGAAACGTTAAAGAAATTTGATGATCTACGAAATATATAAGAAAGATGAT 292
Db 100161 TTTTGAAGAACTTTAAAGAAATTTGATGATCTATGAATAATATAAGAAAGATGAT 100220
QY 293 TTAACACGAGAAAGAGCTCTACAGCAGCTTCTCCAGAGAGCCTAATTAATAGTCAAGAA 352
Db 100221 TCACACGAGAAAGAGCTCTACAGCAGCTTCTCCAGAGAGCCTAATTAATAGTCAAGAA 100280
QY 353 TTGGGAGATGAAAAATACAGATTGTTACACAAATGCTCGAATTTGGTGGAAATCGGCA 412
Db 100281 TTGGGAGATGAAAAATACAGATTGTTACACAAATGCTCGAATTTGGTGGAAATCGGCA 100340
QY 413 AGACAAATGGATACACTACAGTGTTCAGAGATGTTTCCAGATGCTGCTGAAAGTGAACGAGCTCA 472
Db 100341 AGACAAATGGATGATGATTCACAGTGTTCAGAGATGCTGCTGAAAGTGAACGAGCTCA 100400
QY 473 GATAAGCAAGATGGATTCAGCAAGTTCAGCAAGATGTTTCAAGAGAGCCCGCAGGCGAG 532
Db 100401 GACAAATCAAGATGGATTCAGCAAGTTCAGCAAGATGTTTCAAGAGAGCCCGCAGGCGAG 100460
QY 533 CGGACAGTGAAGCGGTGATTTATGTCACATGGCAAAATGGGATGGAAGACATGTGATGAT 592
Db 100461 CGGACAGTGAAGCGGTGATTTATGTCACATGGCAAAATGGGATGGAAGACATGTGATGAT 100520
QY 593 CAGCCACCTTAAAGAAAGAAATCCAAAGTCAGCAAGAAAGAAAGAAAGCTCCAGGCCAAG 652
Db 100521 CAGCCACCTTAAAGAAAGAAATCCAAAGTCAGCAAGAAAGAAAGCTCCAGGCCAAG 100580
QY 653 CAGGAAGGGAAGCTTACCTGTTGAGTTGCAATAGATCCCTTAATGAACCTACATCTGTC 712
Db 100581 CAGGAAGGGAAGCTTACCTGTTGAGTTGCAATAGATCCCTTAATGAACCTACATCTGTC 100640
QY 713 TTATGCAACCAAGTGTCTTATGGGAGATGATGATGATGATGATGATGATGATGATGAT 772
Db 100641 TTATGTAACCAAGTGTCTTATGGGAGATGATGATGATGATGATGATGATGATGAT 100700
QY 773 GAATGTTTTCACCTTTTCATGCTTTTCACCTTACCTTACCTTACCTTACCTTACCTTACCT 832
Db 100701 GAGTGGTTTTCACCTTTTCATGCTTTTCACCTTACCTTACCTTACCTTACCTTACCT 100760
QY 833 CCAAGTGCAGGGGAGATAATGAGAAACAAATGGACAAAGTACTGAAAGACAAAGAAAG 892
Db 100761 CCAAGTGCAGGGGAGATAATGAGAAACAAATGGACAAAGTACTGAAAGACAAAGAAAG 100820
QY 893 GATAGAAATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 952
Db 100821 GATAGAAATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 100879
QY 953 AATTCGTTTTCCTTTCAGAAATGTTTTCAGAAATGTTTTCAGAAATGTTTTCAGAAATGTTT 1012
Db 100880 AATTCGTTTTCCTTTCAGAAATGTTTTCAGAAATGTTTTCAGAAATGTTTTCAGAAATGTTT 100939
QY 1013 AATCATTAGTATTAATGTTGTTTAAAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1054
Db 100940 AATCATTAGTATTAATGTTGTTTAAAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 100981
```

AC114762 42029 bp DNA linear HTG 25-MAR-2002  
Homo sapiens chromosome 4 clone RP11-366M5, \*\*\* SEQUENCING IN  
PROGRESS \*\*\*, 26 unordered pieces.

AC114762  
HTG: HTGS\_PHASE1.  
Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 42029)  
Waterston,R.H.  
The sequence of Homo sapiens clone  
Unpublished  
2 (bases 1 to 42029)  
Waterston,R.H.  
Direct Submission  
Submitted (11-MAR-2002) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
3 (bases 1 to 42029)  
Waterston,R.H.  
Direct Submission  
Submitted (25-MAR-2002) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
On Mar 25, 2002 this sequence version replaced gi:19339148.

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu/gsc/index.shtml  
Contact: submissions@wustl.wustl.edu  
----- Project Information -----  
Center project name: H\_NH0366M05

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 26 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1408: contig of 1408 bp in length  
\* 1409: gap of unknown length  
\* 1509: contig of 1220 bp in length  
\* 2729: gap of unknown length  
\* 2829: contig of 1074 bp in length  
\* 3903: gap of unknown length  
\* 4002: contig of 1060 bp in length  
\* 5063: gap of unknown length  
\* 6295: contig of 1132 bp in length  
\* 6394: gap of unknown length  
\* 6395: contig of 1104 bp in length  
\* 7498: gap of unknown length  
\* 7599: contig of 1196 bp in length  
\* 8795: gap of unknown length  
\* 8895: contig of 1475 bp in length  
\* 10370: gap of unknown length  
\* 10470: contig of 1179 bp in length  
\* 11749: gap of unknown length  
\* 11749: contig of 1453 bp in length  
\* 13202: gap of unknown length  
\* 13302: contig of 1200 bp in length  
\* 14501: gap of unknown length  
\* 14602: contig of 1291 bp in length  
\* 15992: gap of unknown length  
\* 15993: contig of 1531 bp in length  
\* 17623: gap of unknown length  
\* 17624: contig of 1116 bp in length

```

* 18740 18839: gap of unknown length
* 18840 20023: contig of 1184 bp in length
* 20024 20123: gap of unknown length
* 20124 21315: contig of 1192 bp in length
* 21316 21415: gap of unknown length
* 21416 22823: contig of 1408 bp in length
* 22824 22923: gap of unknown length
* 22924 24116: contig of 1193 bp in length
* 24117 24216: gap of unknown length
* 24217 26155: contig of 1939 bp in length
* 26156 26255: gap of unknown length
* 26256 27262: contig of 1007 bp in length
* 27263 27362: gap of unknown length
* 27363 29456: contig of 2094 bp in length
* 29457 29556: gap of unknown length
* 29557 32466: contig of 2910 bp in length
* 32467 32566: gap of unknown length
* 32567 34198: contig of 1632 bp in length
* 34199 34299: gap of unknown length
* 34299 36268: contig of 1970 bp in length
* 36269 36368: gap of unknown length
* 36369 38690: contig of 2322 bp in length
* 38691 38790: gap of unknown length
* 38791 42029: contig of 3239 bp in length.

```

## FEATURES

source

```

1..42029
  Location/Qualifiers
    organism="Homo sapiens"
    db_xref="taxon:9606"
    chromosome="4"
    clone="RP11-366M5"
  1..1408
    /note="assembly_name:Contig46"
  1509..2728
    /note="assembly_name:Contig48"
  2829..3902
    /note="assembly_name:Contig49"
  4003..5062
    /note="assembly_name:Contig50"
  5163..6294
    /note="assembly_name:Contig52"
  6395..7498
    /note="assembly_name:Contig53"
  7599..8794
    /note="assembly_name:Contig55"
  8895..10369
    /note="assembly_name:Contig57"
  10470..11648
    /note="assembly_name:Contig58"
  11749..13201
    /note="assembly_name:Contig60"
  13302..14501
    /note="assembly_name:Contig61"
  14602..15892
    /note="assembly_name:Contig62"
  15993..17523
    /note="assembly_name:Contig63"
  17624..18739
    /note="assembly_name:Contig64"
  18840..20023
    /note="assembly_name:Contig65"
  20124..21315
    /note="assembly_name:Contig66"
  21416..22823
    /note="assembly_name:Contig68"
  22924..24116
    /note="assembly_name:Contig69"
  24217..26155
    /note="assembly_name:Contig70"
  26256..27262
    /note="assembly_name:Contig71"
  27363..29456
    /note="assembly_name:Contig73"
  29557..32466

```

```

misc_feature      /note="assembly_name:Contig74"
32567..34198
misc_feature      /note="assembly_name:Contig75"
34299..36268
misc_feature      /note="assembly_name:Contig76"
36369..38690
misc_feature      /note="assembly_name:Contig77"
38791..42029
/note="assembly_name:Contig78"
BASE COUNT      11707 a 8111 c 8009 g 11615 t 2587 others
ORIGIN
Query Match      53.2%; Score 574.6; DB 2; Length 42029;
Best Local Similarity 84.7%; Pred. NO. 4.8e-113;
Matches 577; Conservative 0; Mismatches 104; Indels 0; Gaps 0;
Qy 233 TATCAAGAAAGCTTTAAAGGAAATTTGATGATCTCTACGAAAAATATAAGAAAGAGATGAT 292
Db 15179 TTTTGTAGAAAGCTTTAAAGGAAATTTGATGATCTCTACGAAAAATATAAGAAAGAGATGAT 15120
Qy 293 TTAACCCAGAGAAAGCTCTACAGAGCTTCTCCAGAGAGCAGCTAAATTAATAGTCAAGAA 352
Db 15119 TTAACCCAGAGAAAGCTCTACAGAGCTTCTCCAGAGAGCAGCTAAATTAATAGTCAAGAA 15060
Qy 353 TTGGGAGATGAAAAAATACAGATTGTTTACACAAATGCTCGAATTTGGTGGAAAAATCGGGCA 412
Db 15059 TTGGGAGATGAAAAAATACAGATTGTTTACACAAATGCTCGAATTTGGTGGAAAAATCGGGCA 15000
Qy 413 AGACAAATGGAGTTTACACTACAGTGTTCACAAATCTCTGTTGAAAGTGAACGAGCTTCA 472
Db 14999 AGACAAATGGAGTTTACACTACAGTGTTCACAAATCTCTGTTGAAAGTGAACGAGCTTCA 14940
Qy 473 GATAAGCAAAAGATGGATTCCAGCCCAACCAAGAAAGATCTTCAAGAAAGACCCCGCAGGCG 532
Db 14939 GATAAGCAAAAGATGGATTCCAGCCCAACCAAGAAAGATCTTCAAGAAAGACCCCGCAGGCG 14880
Qy 533 CGGACCAGTGAAGCCCGTGATTTATGTACATGCAATGGGATTTGAAGACTGTGATGAT 592
Db 14879 CGGACCAGTGAAGCCCGTGATTTATGTACATGCAATGGGATTTGAAGACTGTGATGAT 14820
Qy 593 CAGCCACCTAAGAAAGAAATCCAGTCACAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 652
Db 14819 CAGCCACCTAAGAAAGAAATCCAGTCACAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 14760
Qy 653 CAGAAAGGGAGAGCTTCACCTGTTGAGTTTGCAATAGATCCTAATGAACCTACATACCTGC 712
Db 14759 CAGAAAGGGAGAGCTTCACCTGTTGAGTTTGCAATAGATCCTAATGAACCTACATACCTGC 14700
Qy 713 TTATGCAACCAAGTGTCTTATGGGAGATGATAGATGATGATGATGATGATGATGATGATGAT 772
Db 14699 TTATGCAACCAAGTGTCTTATGGGAGATGATAGATGATGATGATGATGATGATGATGATGAT 14640
Qy 773 GAATGGTTTCACTTTTCATGTTTTCATCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 832
Db 14639 GAATGGTTTCACTTTTCATGTTTTCATCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 14580
Qy 833 CCAAGAGTGCAGGGGAGATAATGAGAAAAACAATGAGAAAAACAATGAGAAAAACAATGAGAAAAACA 892
Db 14579 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 14520
Qy 893 GATAGAAGATCGAGGTAGTAA 913
Db 14519 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
LOCUS
AY014017
DEFINITION Xenopus laevis p3INGL mRNA, complete cds.
ACCESSION AY014017
VERSION AY014017.1 GI:17864714
KEYWORDS
Xenopus laevis.

```

## RESULT 11

AY014017

LOCUS

AY014017

DEFINITION

Xenopus laevis

p3INGL mRNA, complete cds.

ACCESSION

AY014017

VERSION

AY014017.1

GI:17864714

KEYWORDS

Xenopus laevis.

AY014017

1094 bp

mRNA

linear

VRT 16-DEC-2001

RESULT 12	AC129658	188087 bp	DNA	linear	HTG 31-JUL-2002
LOCUS	AC129658				
DEFINITION	Rattus norvegicus clone CH230-6112, *** SEQUENCING IN PROGRESS ***;				
	69 unordered pieces.				
ACCESSION	AC129658				
VERSION	AC129658.1	GI:22024520			
KEYWORDS	HTG; HTGS_PHASE1.				
SOURCE	Norway rat.				
ORGANISM	Rattus norvegicus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;				
	Rattus.				
REFERENCE	1 (bases 1 to 188087)				
AUTHORS	Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Anarattung,H.C., Are,J.R., Ayele,M., Banks,T., Barbaria,J., Benton,J., Bimager,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Buahy,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carton,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinsh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B., Homi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,				

```

Kratochvic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
Sodergren,E., Sonaika,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G., and Gibbs,R.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GBVT
Center clone name: CH230-6112
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 137386 bases at least Q40
Consensus quality: 145233 bases at least Q30
Consensus quality: 150556 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: this is a 'working draft' sequence; it currently
* consists of 69 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 1002: contig of 1002 bp in length
* 1003 1102: gap of unknown length
* 1103 2115: contig of 1013 bp in length
* 2116 2215: gap of unknown length
* 2216 3536: contig of 1321 bp in length
* 3537 3636: gap of unknown length
* 3637 4943: contig of 1307 bp in length
* 4944 5043: gap of unknown length
* 5044 6443: contig of 1400 bp in length
* 6444 6543: gap of unknown length
* 6544 7965: contig of 1422 bp in length
* 7966 8065: gap of unknown length
* 8066 9338: contig of 1273 bp in length
* 9339 9439: gap of unknown length
* 9439 11165: contig of 1727 bp in length
* 11166 11265: gap of unknown length
* 11266 12838: contig of 1573 bp in length
* 12839 12938: gap of unknown length
* 12939 14397: contig of 1459 bp in length
*
14398 14497: gap of unknown length
14498 15552: contig of 1055 bp in length
15553 17205: contig of unknown length
17206 17305: gap of unknown length
17306 18474: contig of 1169 bp in length
18475 20060: contig of unknown length
20061 20160: gap of unknown length
20161 21232: contig of 1072 bp in length
21233 21332: gap of unknown length
21333 23527: contig of 2195 bp in length
23528 23627: gap of unknown length
23629 25397: contig of 1770 bp in length
25398 25497: gap of unknown length
25498 26707: contig of 1210 bp in length
26708 26807: gap of unknown length
26808 28142: contig of 1335 bp in length
28143 28242: gap of unknown length
28243 29532: contig of 1290 bp in length
29533 29632: gap of unknown length
29633 32229: contig of 2597 bp in length
32230 32329: gap of unknown length
32330 34198: contig of 1869 bp in length
34199 34298: gap of unknown length
34299 35844: contig of 1446 bp in length
35845 37406: contig of 1562 bp in length
37407 37506: gap of unknown length
37507 39608: contig of 2102 bp in length
39609 39708: gap of unknown length
39709 41834: contig of 2126 bp in length
41835 41934: gap of unknown length
41935 43565: contig of 1631 bp in length
43566 43665: gap of unknown length
43666 45315: contig of 1650 bp in length
45316 45416: gap of unknown length
45417 47171: contig of 1755 bp in length
47171 49186: contig of 1916 bp in length
49187 49287: gap of unknown length
49287 50868: contig of 1581 bp in length
50868 53214: contig of 2247 bp in length
53215 53315: gap of unknown length
53315 55155: contig of 1841 bp in length
55156 55255: gap of unknown length
55256 57107: contig of 1852 bp in length
57108 59981: contig of 2774 bp in length
59982 60081: gap of unknown length
60082 62198: contig of 2117 bp in length
62199 62298: gap of unknown length
62299 63826: contig of 1528 bp in length
63827 63926: gap of unknown length
63927 65474: contig of 1548 bp in length
65475 65574: gap of unknown length
65575 68299: contig of 2725 bp in length
68300 68399: gap of unknown length
68400 69799: contig of 1400 bp in length
69800 72245: gap of unknown length
69900 72246 72345: gap of unknown length
72346 74214: contig of 1869 bp in length
74215 74314: gap of unknown length
74315 76029: contig of 1715 bp in length
76030 76129: gap of unknown length
76130 79736: contig of 3607 bp in length
79737 79836: gap of unknown length
79837 81351: contig of 1515 bp in length
81352 81451: gap of unknown length
81452 83987: contig of 2536 bp in length
83988 84087: gap of unknown length

```

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

```

* 84088 86246: contig of 2159 bp in length
* 86247 86346: gap of unknown length
* 86347 89935: contig of 3589 bp in length
* 89936 90036: gap of unknown length
* 90036 92953: contig of 2918 bp in length
* 92954 93053: gap of unknown length
* 93054 95875: contig of 2822 bp in length
* 95876 95975: gap of unknown length
* 95976 98560: contig of 2585 bp in length
* 98561 98661: gap of unknown length
* 98661 101078: contig of 2418 bp in length
* 101079 101178: gap of unknown length
* 101179 103586: contig of 2408 bp in length
* 103587 103686: gap of unknown length
* 103687 106731: contig of 3045 bp in length
* 106732 106831: gap of unknown length
* 106832 109375: contig of 2544 bp in length
* 109376 109475: gap of unknown length
* 109476 112710: contig of 3235 bp in length
* 112711 112811: gap of unknown length
* 112811 116078: contig of 3267 bp in length
* 116078 116177: gap of unknown length

Query Match      25.3%; Score 273.6; DB 2; Length 188087;
Best Local Similarity 92.3%; Pred. No. 4.4e-48;
Matches 310; Conservative 0; Mismatches 24; Indels 2; Gaps 2;

QY 720 ACCAAGTGTCTTAT-GGGGAGATGATAGGATGACAAATGACAGTGTCCAAATTTGAATGG 778
Db 146506 ATCAAGTGTCTTATCCGGGAGATGATAGGCTGTGACAAATGATAGGCTCCCATTTAGTGG 146505

QY 779 TTTCACCTTTTCATGTTTTCACCTTACCTATATAACCAAGGGGAAATGGTATGCCCAAAG 838
Db 146566 TTTCACCTTTTCTGCTTTTCACTTCACTATATAACCAAGGGGAAATGGTATGCCCAAAG 146625

QY 839 TGCAGGGGAGATAATCAGAAACAAATGGACAAAGTACTGAAAGACAAAGGATAGA 898
Db 146626 TGCAGGGGAGACAGTGGAAACCAATGGACAAAGTACTGAAAGACAAAGGAGAGA 146685

QY 899 AGATCGAGGTAGTAAAGGCCATCCACATTTTAAAGGGTATTTGTCTTTATATATATTCG 958
Db 146686 AGAGCGAGGTAGTAAAGGCCATCC-GGTTTAAAGGGTCTTTGTCTTTATATATTCG 146744

QY 959 TTTCCTTTTCAGAAATGTTTATAGGTAATGATGATGATGATGATGATGATGATGATGAT 1018
Db 146745 TTTCCTTTTCAGAAATGTTTATAGGTAATGATGATGATGATGATGATGATGATGATGAT 146804

QY 1019 TAGTATTAATGTTGTTTAAAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1054
Db 146805 TAGTATTAATGTTGTTTAAAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 146840

RESULT 13
AF149820
LOCUS      AF149820
DEFINITION Mus musculus p33ING1 protein mRNA, linear
ACCESSION AF149820
VERSION    AF149820.1 GI:6456561
KEYWORDS
SOURCE     Mus musculus.
ORGANISM   Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 840)
Rancourt,D. and Garikavtsev,I.
Structural organization and expression pattern of the murine INCL
gene
Unpublished
2 (bases 1 to 840)
Garikavtsev,I.V.
Direct Submission
Submitted (11-MAY-1999) Genome Therapeutics Corp., 100 Beaver St.,
Waltham, MA 02453, USA

```

```

FEATURES             Location/Qualifiers
     source            1..840
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     CDS               1..840
                     /codon_start=1
                     /product="p33ING1 protein"
                     /protein_id="AAF09183.1"
                     /db_xref="GI:6456562"
                     /translation="MLSPANGQIHLVNYVEDYLSIESLPFDLQNVSLMRIDAKY
                     QEILKELDDIYEKFKRETGTQRRVLRHICQIALIRISQELGDEKIQIVSQMVELVNR
                     SROVDSHDFEAFHQDISDGTGSGKAGQKSEAITQADKPNKRSRRQNNRE
                     NASNNHDDITSGTPKPKKAKTSKKKRSKABREASPADPDPIDNPPTVCLCNQV
                     SYGEMIGCDNDECPIEWFHFSCVGLNHHKPKGNWYCPKCRGESEKTMKALEKSKKRA
                     YNR"
BASE COUNT      250 a   216 c   251 g   123 t
ORIGIN
Query Match      23.8%; Score 256.8; DB 10; Length 840;
Best Local Similarity 60.6%; Pred. No. 5.5e-45;
Matches 492; Conservative 0; Mismatches 287; Indels 33; Gaps 3;

QY 116 ACCGGGAGCGGAGCGGCTGCTACCTGCTAGTGTGACAGACTACCTTCTAGTGTGCGGAG 175
Db 16 AACGGGAGCAGATCCACCTGGTGA---ACTATGTGGAGGATTACCTGTGACCAATCGAG 72

QY 176 TCGCTGCCCCACGACATGCAGAGAACGTGTCTGTCTGCGAGAGCTGCACACCAATAT 235
Db 73 TCACCTGCTTTTCGACCTGCAGAGAACGTCTCTGCTGATCGGGAGATCCACGCCAATAC 132

QY 236 CAAGAACGTTTAAAGAAATTTGATGATGCTACGAAAAATATAAGAAAGAGATATTTA 295
Db 133 CAAGAGATCTCAAGAGAGCTGGACGACTACTATGAGAAGTTCAAAACGGGAGACAGACGGC 192

QY 296 RACCAGAAAGACGCTACAGCAGCTTCTCCAGAGAGCACTAATTAATAGTCAAGATTG 355
Db 193 ACCCAGAGCGCGGGTATGCTACTGCTATCCAGAGAGGGCCCTGATCCGACGACGAGCTA 252

QY 356 GGAGATGAAAAAATACAGATTGTTACACAAATGCTCGAATTTGTTGMAAATCGGGCAAGA 415
Db 253 GCGATGAGAACATCCAGATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 312

QY 416 CAAATGGAGTTACATCACAGTGTGTTTCCAAAGATCTGTTGTTGTTGTTGTTGTTGTTG 458
Db 313 CAGGTGGACAGTCAGTGGAGCTCTTCGAAGCACACACAGGACATCAGTGACGCGACTGCT 372

QY 459 -----GTGAAGAGCCTCAGATAAGCAAGATGATGTTCCAGCAACCAACCAAGATCT 511
Db 373 GGCAGCGGCAAGCGGCGGAGCAAGTGAAGAGTGAAGGCTCCTCAGCAGCAGATAGAG 432

QY 512 TCAAGAA-----GACCCCGCAGGCGGAGGACCGAGTGAAGCGGTGATTATGTACATG 565
Db 433 CCGAATAACAACGGTTCCAGGAGGCGAGCAACATGAGATCGAGAGAACCGTCTGAAT 492

QY 566 GCAATGGATTGAAGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 625
Db 493 AATCAGCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 552

QY 626 AAGAAAGAGACGCTCCAAAGCCAGCAGGAGGAGGAGCTTCACTGTTGAGTTTGA 685
Db 553 AAGAAGAGAAACGGTCCAAAGCCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 612

QY 686 ATAGATCTCTAATGAACCTACATACCTATGCTATGCAACCAAGTCTCTTATGGGAGATGATA 745
Db 613 ATCGACCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 672

QY 746 GGATGTGACAAATGAACAGTGTCCAAATGAATGGTTTCACTTTTCAATGTTTCACTTACC 805
Db 673 GGCTGTGACACGACGAAATGCCCATCGAGTGTGTTCCACTTCTCTCTGCTGGGGCTCAAC 732

QY 806 TATAAACCAAGGGGAAATGTTATGTTCCCAAGTCCAGGGGAGGAGATTAATGAGAAACAATG 865
Db 733 CATAAACCAAGGGGAAATGTTATGTTCCCAAGTCCAGGGGAGGAGGAGGAGGAGGAGGAGGAG 792

```



QY	866	GACAAAGTACTGAAAACACAAAAAGATAG	897
Dd	793	GACAAGCCCTTGAGAAGTCCAAGAAGAGAG	824
RESULT 14			
LOCUS	BC016573	1584 bp	mRNA linear ROD 07-AUG-2002
DEFINITION	Mus musculus, inhibitor of growth family, member 1, clone MGC:27881 IMAGE:3496013, mRNA, complete cds.		
ACCESSION	BC016573		
VERSION	BC016573.1	GI:16741521	
KEYWORDS	MGC.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 1584)		
TITLE	Straussberg,R.		
JOURNAL	Direct Submission		
REMARK	Submitted (31-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
COMMENT	NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cypabs@mail.nih.gov Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: http://www.hgsc.bcm.tmc.edu/cdna/ Contact: amg@bcm.tmc.edu Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.		
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAK Plate: 35 Row: i Column: 6.			
FEATURES	Location/Qualifiers		
source	1..1584 /organism="Mus musculus" /db_xref="taxon:10090" /map="C57BL/6J" /clone="MGC:27881 IMAGE:3496013" /tissue_type="mammary tumor. WAP-TGF alpha model. 7 months old, gross tissue." /clone_lib="NCI-CGAP_Mam5" /lab_host="DH10B" /notes="Vector: pCMV-SPORT6" 389..1228 /codon_start=1 /product="inhibitor of growth family, member 1" /protein_id="AAH16573.1" /db_xref="GI:16741522" /db_xref="LocusID:26356" /translation="MLSPANGQIHLNVYEDYLDSIESLPDLQRNYSMLREIDAKY OEILKELDYEFKRETGTOKRRLVHCIOALIRSOELGDKEIQIVSQMVLEVNRS QWDSHVLFEEAHQDISGTGGSGKAGDQSKSEAITQADKPNNKRORRNENRE NASNNHHDDITTSQTPEKKAKTSKKRAERASPADLPIDPNEPYICLNQV SYGEMIGDCNDCEPIEFHFSCVLGNHPKGWKYCPKCRGESEKTMDALEKSRRKA YNR"		
CDS	384 a 422 c 517 g 261 t		
BASE COUNT			
ORIGIN			
Query Match	23.5%	Score 253.6;	DB 10; Length 1584;
Best Local Similarity	60.3%	Pred. No. 3.le-44;	
Matches 490; Conservative	0;	Mismatches 289;	Indels 33; Gaps 3;

Search completed: June 20, 2003, 02:14:51  
Job time : 2830 secs

Search completed: June 20, 2003, 02:14:51  
Job time : 2830 secs